COMPRESSED FLUORESCENCE LIFETIME IMAGING MICROSCOPE

BY

YAYAO MA

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Adviser:

Assistant Professor Liang Gao

Abstract

Fluorescence lifetime imaging microscopy (FLIM) has been extensively applied in wide biomedical applications from single cell studies to medical diagnosis. While the state-of-the-art FLIM techniques are highly sensitive to biological dynamics such as fluorescence resonance emission transfer (FRET), they depend on repetitive measurements and the slow frame rate limits the study of ultrafast biological dynamics. This thesis reports the world's fastest high-resolution FLIM at 100 frames per second (fps). This system, referred to as compressed FLIM, can capture a widefield lifetime image of 500 × 450 pixels within a single camera snapshot and operate real-time two-dimensional (2D) FLIM at 100 fps. By combining compressed ultrafast photography (CUP) with FLIM and adopting dual-camera detection, compressed FLIM is demonstrated for both fluorescent biological samples lifetime imaging as well as 75 fps image acquisition of fluorescent beads diffusion dynamics. Compressed FLIM holds great promise for quantitative live cell imaging applications. Compressed FLIM also has great potential in high-speed imaging of fluorescence lifetime in transient biological events such as FRET and may further extend quantitative live cell imaging such as monitoring neural spiking.

Acknowledgments

I would like to thank my advisor, Professor Liang Gao, for his continuous support, encouragement, and valuable insights throughout the learning process of this thesis work. It has been a truly enjoyable experience to work in the iOptics lab.

I also would like to thank Dr. Xiaohua Feng for his patience and generosity in sharing his immense expertise and valuable life lessons with me. Special thanks to all of my friends and colleagues in the iOptics lab for being supportive and caring and making this lab the best place to work.

Finally, I would like to thank my family for always being supportive and understanding wherever I go and whatever I do. I would like to thank Jay Chou and Yoona Lim for encouraging me during all the ups and downs.

Table of Contents

1. Background

1.1. Fluorescence process

Fluorescence of fluorophores is a light-matter interaction process and contains three events: excitation, vibration relaxation (internal conversion) and fluorescence emission. A Jablonski energy diagram illustrates the process and the electronic transitions between ground states and excited states as is shown in **Figure 1.1**. Photons are absorbed by the molecules whose energy matches the energy difference between the discrete energy levels. These energy levels define the absorption and the emission band. *Excitation* refers to the absorption of a photon and can elevate an electron of a molecule from the ground state to an excited state. Excitation is caused by either mechanical or chemical mechanisms in nature and the energy source of the excitation can be a laser or a mercury lamp in fluorescence or chemical reactions in chemiluminescence. The fluorophores then transit to the excited electronic states. Excitation by a photon occurs in femtoseconds. Vibrational relaxation to the lowest vibrational energy level in the excited state lasts for picoseconds. The initial excited state is partially dissipated and transits to relaxed singlet excited states from which fluorescence emission results. Fluorescence emission lingers over a finite time and molecules return to the ground state. In addition to fluorescence emission, the excited states can be depopulated by other processes such as collisional quenching, fluorescence resonance energy transfer (FRET) and intersystem crossing. The ratio of the number of photons emitted to the number of photons absorbed is referred to as the quantum yield of fluorescence which measures the relative extent among these processes. Energy dissipation originates from the initial vibrational relaxations and the subsequent inclined fluorescence transitions, which derive from excited-state reactions, complex formations and FRET [1], to higher vibrational energy levels of the ground state.

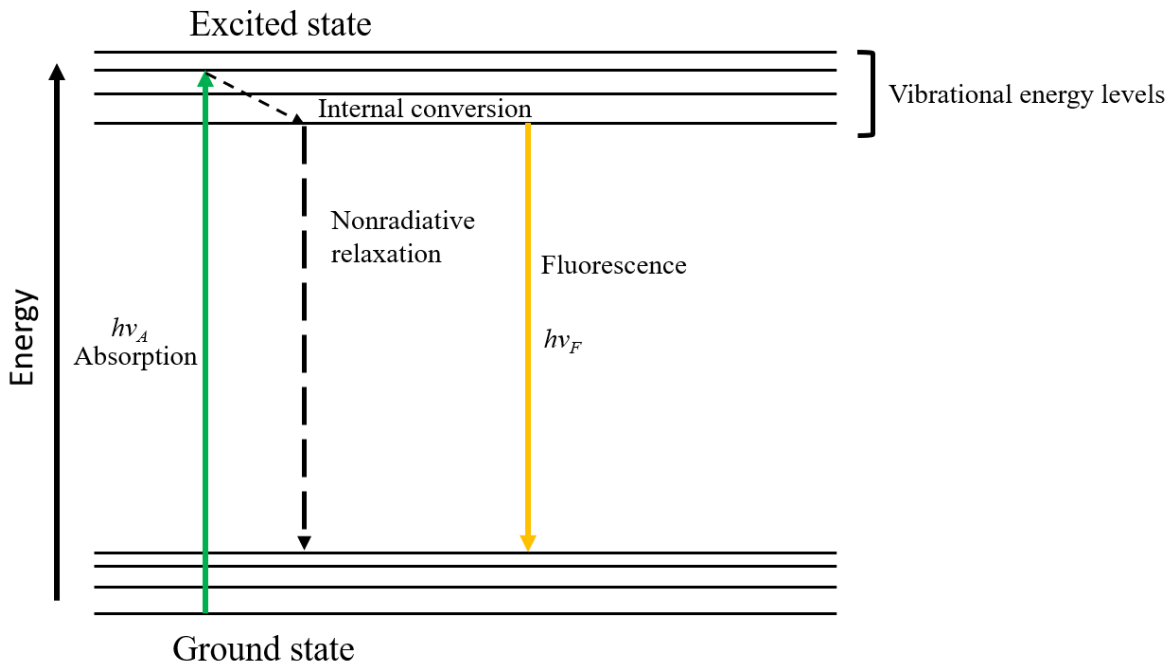

Figure 1.1 Jablonski energy diagram showing excitation and various possible relaxation mechanisms. Each *hv* denotes the photon energy, where subscripts *A*, *F* denote absorption and fluorescence.

Due to energy dissipation, the energy of emitted photons is lower than of the excited photons and the energy difference represented by (hv_A–hv_F) defines the Stokes shift. The Stokes shift results in the red-shifted fluorescence emission relative to the excitation light. Emission wavelengths are independent of the excitation wavelength because of the vibrational relaxation to intermediate excited states. Fluorescence emission is incoherent because of the uncertain delays in the vibrational relaxations. Fluorescence emission spectrum generally is in mirror symmetry to the absorption spectrum. The characteristics of these spectra can be used to distinguish different fluorophores when they are simultaneously detected. Intensity of the excitation at the fluorescence excitation spectrum can affect fluorescence emission intensity. A single fluorophore can produce thousands of detectable photons given that the same fluorophore can be repeatedly excited. *Photobleaching* refers to the phenomenon that one fluorophore is irreversibly destroyed and is permanently unable to fluorescence [2].

Fluorescence detection systems typically consist of four parts: 1) an excitation source, 2) filter sets, 3) a detector, 4) a fluorophore. These four parts should be compatible for the fluorescence detection fidelity. Fluorescence microscopes resolve fluorescence spatially for microscopic fluorescent samples. Fluorescence scanners resolve fluorescence spatially for macroscopic objects.

Flow cytometers measure fluorescence characteristics in a large flowing particle stream to distinguish and isolate particles. Other fluorescence instrumentations include DNA sequencers and microfluidic devices. Different types of fluorescence instruments have different requirements on the fluorophores and measurements. For example, photobleaching should be avoided in fluorescence microscopy. However, it is not a significant problem in flow cytometry because of high flow speed and short exposure time under the excitation beam of the particles.

Fluorescence intensity depends on the same parameters as the absorbance and the fluorescence quantum yield of the target fluorophores, namely, on the excitation intensity and fluorescence collection efficiency of the detection instrument. In dilute solutions, fluorescence intensity is almost linearly proportional to the above parameters. In dense solutions, this relationship can be further distorted by self-absorption and the inner-filter effect. Background signals, which can result from autofluorescence or reagent background, seriously distort fluorescence detection sensitivity. Autofluorescence can be minimized by either appropriate filter sets or target fluorophores that absorb the autofluorescence. Reducing the detection bandwidth can increase the detection resolution but compromises the overall fluorescence intensity. Using fluorophores with > 500 nm excitation capability can mostly minimize the autofluorescence from cells and tissues and other biological samples. Moreover, longer emission wavelengths reduce the light scattering by the dense media as well and contribute to greater penetration of the excitation light.

1.2. Fluorescence lifetime

Lifetime is defined as the average time that an excited molecule spends in the excited state before returning to the ground state. Lifetime decides the time available for the fluorophores to interact with their environment and the information revealed from their emission. Γ defines the emissive rate of the fluorophore. k_{nr} refers to the rate of non-radiative decay to the ground state. The quantum yield Q determines the part of the fluorophores that decay through emission:

$$Q = \frac{\Gamma}{\Gamma + k_{nr}}. \tag{1.1}$$

Lifetime τ is

$$\tau = \frac{1}{\Gamma + k_{nr}}. \tag{1.2}$$

Fluorescence emission is a scholastic process and for a single exponential decay, 63% of the molecules have decayed prior to $t = \tau$ [3].

1.3. Fluorescence probes

In order to detect biological molecules such as proteins, polypeptides, antibodies and DNA/RNA and intracellular biological structures such as nuclei, mitochondria and collagen, a fluorescent probe is attached chemically or biologically to these molecules and structures. Generally, fluorescent probes use the fluorophores which are highly stable and sensitive and selectively bind to one specific region or functional group of the target molecular. A fluorophore can emit fluorescence upon light excitation. It absorbs photons at a specific energy level and re-emit photons at a lower energy level and longer wavelength. Maximum excitation and emission wavelength define the peak in the excitation and emission spectra, respectively. During fluorescence, the fluorophore exhibits a time-resolved decay pattern that can be easily characterized with a decay rate of $1/\tau$ and lifetime is τ equal to the time taken to decay to $1/e$ (≈ 0.368) of the original emission level and the decay process is mono-exponential in most cases. Lifetime of the fluorophores is independent of concentration, sample absorption, sample thickness, excitation intensity and photobleaching. So, it is highly robust and suitable to characterize a fluorophore and the correspondingly labelled protein. In the meanwhile, the fluorescence lifetime is relevant to environmental factors such as solution pH, ion concentration, molecular binding and proximity of energy acceptors. Consequently, the lifetime can be applied to functional imaging and reflects the interaction between labelled cellular structures and the environment.

Fluorophores can be characterized and differentiated by the emission spectrum and fluorescence lifetime practically. In order to differentiate intracellular structures, multiple fluorophores are used to label different subcellular structures such as nuclei and mitochondria. Consequently, the ability to image them simultaneously (multiplexing) and separate them has been critical for understanding a variety of biological processes. Generally, two methods have been applied to achieve multiplexing. The first is the spectrum-based method and it is intensity-based, conventional and widely adopted and separates fluorophores with their emission spectra. The second method is based on the fluorescence lifetime, an identifying characteristic separate from the emission spectrum, which can be used to distinguish fluorophores. This approach relies on the fact that fluorophores exhibit a time-resolved decay pattern that can be easily characterized and is very often mono-exponential with a decay rate of $1/\tau$. Lifetime contrast is independent of fluorescence intensity and emission spectrum and is more robust than intensity-based methods. Fluorescence lifetime imaging microscopy is based on the lifetime contrast.

1.4. Fluorescence lifetime imaging microscopy (FLIM)

Fluorescence lifetime imaging microscopy (FLIM) [4] has been extensively employed in a wide spectrum of biomedical applications, ranging from single cell studies [5] to medical diagnosis [6]. Rather than imaging the time-integrated fluorescent signals, FLIM measures the time-lapse fluorescent decay. Fluorescence lifetime imaging microscopy (FLIM) produces an image based on the differences in the excited state decay rate from a fluorescent sample as illustrated in **Figure 1.2**. Thus, FLIM is a fluorescence imaging technique where the contrast is based on the lifetime of individual fluorophores rather than their emission spectra. As an intrinsic characteristic of a fluorophore and its state, fluorescence lifetime does not depend on concentration, absorption by the sample, sample thickness, photo-bleaching and excitation intensity; it is more robust than intensity-based methods. At the same time, the fluorescence lifetime depends on numerous environmental parameters such as pH, ion or oxygen concentration, molecular binding or the proximity of energy acceptors, making it the technique of choice for functional imaging because lifetime measurements are independent of intensity which can be significantly altered by tissue heterogeneities and depth location. FLIM enables a more quantitative study of molecular effects inside living organisms compared with conventional intensity-based approaches and is widely applied in various fields.

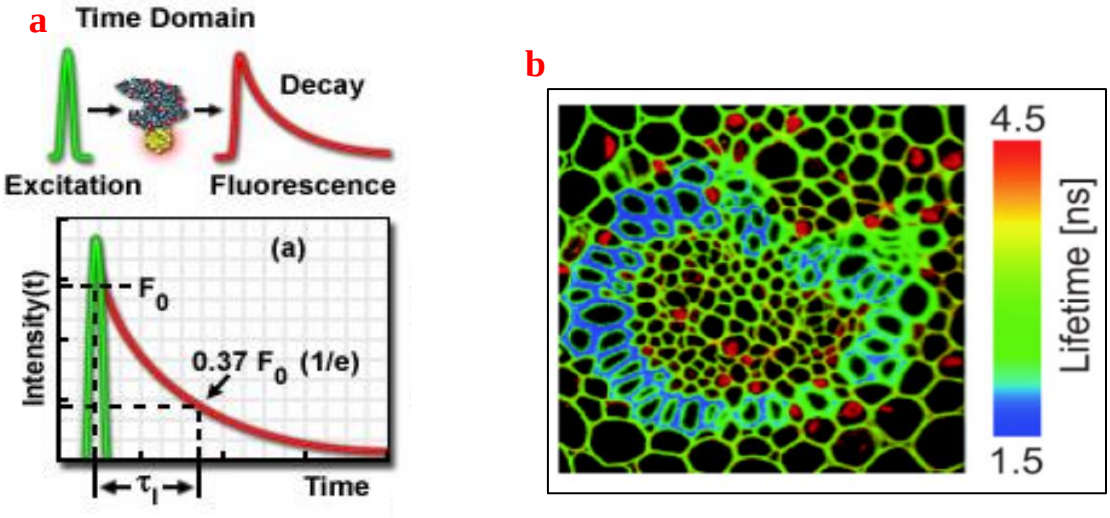

Figure 1.2 FLIM working principle. (a) FLIM measures the time-lapse fluorescent decay. **(b)** FLIM produces an image based on the differences in the excited state decay rate from a fluorescent sample.

There are a variety of technical implementations of FLIM. Based on the detection mechanism, they are generally classified into two categories: time-domain FLIM and frequency-domain FLIM. Time-domain FLIM illuminates the sample with pulsed laser excitation, followed by measuring the fluorescent decay in sequential time channels using an ultrafast detector or detector array. Time-domain FLIM is performed using either scanning-based systems—such as a time-resolved confocal microscope [7] or a streak camera [8], or widefield-based systems—such as a temporally-gated 2D camera or a single photon avalanche diode (SPAD) array. By contrast, frequency-domain FLIM illuminates the sample with high-frequency pulsed or modulated continuous-wave lasers, driving the fluorescence oscillating at the same frequency, however, with a reduced modulation degree and a phase shift due to the fluorophore's lifetime. To measure the modulation and phase of emitted fluorescence, the imager also modulates the gain of detector at the same frequency as the excitation light (homodyne methods) or a frequency slightly different from that of the excitation light (heterodyne methods). The measured data is then analyzed using the phasor approach.

Time-correlated single photon counting (TCSPC) is the most widely used method in time-domain FLIM to acquire fluorescence lifetime. TCSPC is a statistical method and relies on a large number of repetitive laser excitation and delay measurement cycles to accumulate enough photons for the delay histogram. In the case of intensive excitation as shown in **Figure 1.3 a**, many electrons in the ground state of the fluorophore molecules are excited, resulting in multiple photon emission. TCSPC uses extremely low excitation intensity so that one single photon is emitted upon every laser excitation pulse as shown in **Figure 1.3 b**. In TCSPC, the time interval between sample excitation laser pulse and the arrival of the emitted single photon at the detector is repeatedly measured to account for the statistical estimation of the fluorophore's emission decay profile at every laser excitation cycle as shown in **Figure 1.4 a**. A large amount of data can be collected over many cycles. The delay time intervals are constructed into a histogram that summarizes the photon emission occurrence over time upon excitation as shown in **Figure 1.4 b**. The defined 'start' signal is realized by the electronics detecting the laser pulse which restricts the emitted photon counting rates, and the defined 'stop' signal is detected by a single-photon sensitive detector such as a photomultiplier tube (PMT), micro channel plate (MCP), single photon avalanche diode (SPAD) or hybrid PMT. In application, a high number of emitted photons and correspondingly increased pixel dwell time are required to reach sufficient accuracy. However, due to the repetitive

measurement for accuracy and mechanical scanning to the sample for large field of view (FOV), current TCSPC is limited by the slow frame rate.

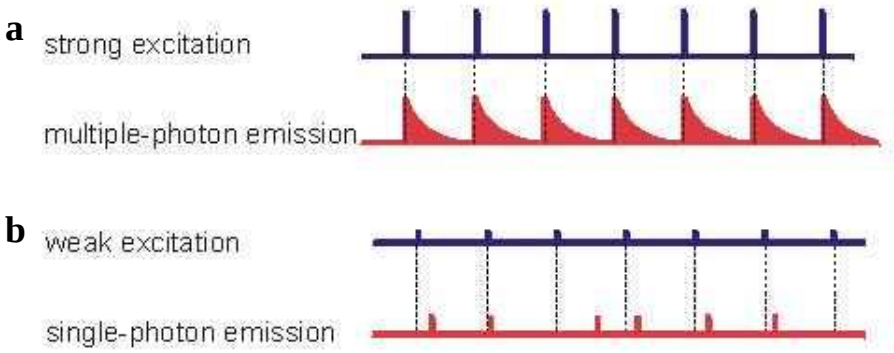

Figure 1.3 Photon emission situations under strong and weak excitation. (a) Many electrons in the ground state of the fluorophore molecules are excited, resulting in multiple photon emission. **(b)** Extremely low excitation intensity is used in TCSPC so that one single photon is emitted upon every laser excitation pulse.

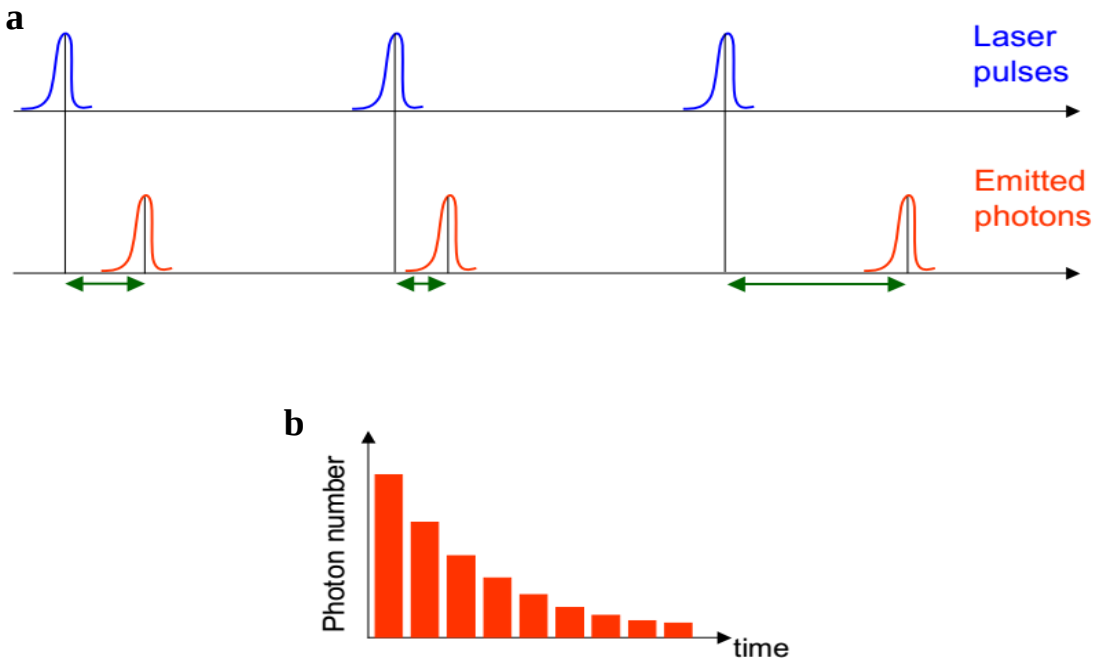

Figure 1.4 TCSPC working principle. **(a)** The time interval between sample excitation laser pulse and the arrival of the emitted single photon at the detector is repeatedly measured to account for the statistical estimation of the fluorophore's emission decay profile at every laser excitation cycle. **(b)** Histogram of delay time intervals in time-resolved fluorescence measurement with TCSPC.

Ultrafast biological dynamics are omniscient in life science due to the microcellular environment and rapid biological reaction. For example, protein folding, catalysis and ligand binding occur within from nanoseconds to milliseconds. Energy and charge transfer, isomerization and bond stretching, and twisting happen within picoseconds. Fluorescence probes are widely used to detect these dynamics. With a reference intensity camera, the intensity-based measurements infer only relative changes, and the results are easily confounded by variations in excitation fluence, signal decay by photobleaching, and background fluorescence, particularly when the imaging is performed in vivo. Compared with intensity-based approaches, FLIM provides a more quantitative assessment on the fluorescent probes because it is insensitive to the aforementioned factors. For example, FLIM has allowed precision detection of FRET in subcellular structures such as dendritic spines in neurons. As another example, confocal FLIM has been used to map the intracellular temperature distribution in live cells. Similarly, in endoscopy, lifetime imaging of tissue autofluorescence has shown great promise in demarcating malignant from normal tissue [9].

1.5. Streak camera

A streak camera is an ultrafast optical imaging instrument that measures a wide range of the light intensity variation with time from a single-shot event. It is employed to measure the pulse duration of laser fusion lasers, free electron lasers and other pulsed lasers and for various time-crucial applications such as time-resolved Raman spectroscopy, fluorescence lifetime measurement, plasma light emission, laser ablation and LIDAR.

Conventionally, a streak camera is a 1D ultrafast imaging device. Its narrow entrance slit (10–50 μm wide) limits the imaging FOV to be a line. It operates by transforming the time profile of light pulse intensity into the spatial profile on a detector through the time-varying deflection of the light. **Figure 1.5** illustrates the operating principle of the streak camera. Assume there are light pulses arriving sequentially at the slit. First, they pass through the slit and lenses and the light image is formed on the photocathode. Here at the photocathode, photons are converted into electrons in the time order and accelerated. Then they pass the sweep electrode sequentially and are deflected vertically at different angles because different sweeping voltages are applied at the different time slots. Then they hit the phosphor screen and are converted back to light, and the image is read out by a digital camera. Finally, in the streak image, the vertical axis records the time information and the horizontal axis captures the space information.

To acquire a 2D high-speed video from the streak camera, one has to scan the imaging FOV along the direction vertical to the streak camera's entrance slit. However, it requires the event to be repetitive at each scanning position.

Temporal resolution of the streak camera is the resolution boundary which differentiates between two consecutive events with regard to the time. In practice, it can be measured by the FWHM (full width at half maximum) of the time profile in the streak image with regard to an incident light pulse whose pulse width can be infinitely close to but not equal to zero [10, 11].

Dynamic range of a streak camera indicates the light intensity range which can be distinguished by the streak camera. In practice, it can be measured by the weakest light pulse which can be actually measured. The ratio between the strongest pulse and the weakest light pulse is taken as the dynamic range. Generally, the dynamic range is in inverse proportion to the temporal resolution.

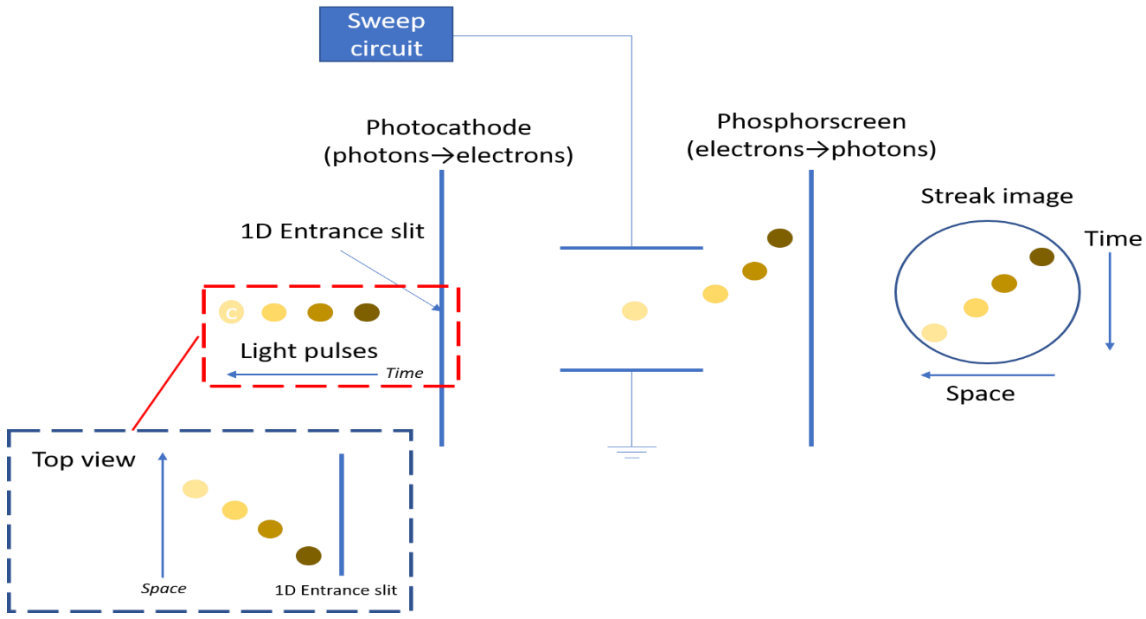

Figure 1.5 Streak camera working principle.

1.6. Compressed ultrafast photography (CUP)

Compressed ultrafast photography (CUP) is the worlds' fastest passive camera [12]. CUP uses conventional imaging optics and a streak camera but takes advantage of the ability of compressed sensing tools to recover images from sparse spatiotemporal data. The approach enables the information in the image to be encoded in the field of view of the streak camera, thereby allowing the system to read out images with full-frame capability at an extremely high speed (up to 100 billion fps). **Figure 1.6** shows the schematic of CUP. **Figure 1.7** shows representative images of CUP at macroscopic scales. The application of compressed sensing in ultrafast imaging has been demonstrated in compressed ultrafast photography (CUP) [12]. Compressed ultrafast photography (CUP) is a new computational ultrafast imaging technology that can capture transient dynamic events at 100 billion frames per second in a single camera exposure with a sequence depth of hundreds of frames. CUP synergistically combines two technologies—the streak camera and compressed sensing (CS) —and can perform 2D ultrafast passive imaging at 100 billion frames per second. Conventionally, a streak camera is a 1D ultrafast imaging device—its narrow entrance slit (10–50 μm wide) limits the imaging FOV to be a line. To acquire a 2D high-speed video, one has to scan the imaging FOV along the direction perpendicular to the streak camera's entrance slit. However, this requirement poses the severe restriction that the event must be repetitive at each scanning position. Unlike other streak-camera-based ultrafast imagers, CUP overcomes this limitation with a fully opened entrance slit onto the streak camera to collect 2D image and uses compressed sensing to reconstruct the 2D image by spatially encoding an image with a random binary pattern using a digital micromirror device (DMD), followed by temporal shearing of the resultant signals using a streak camera. This spatially encoded and temporally sheared image is then detected by a CCD within a single exposure. The image reconstruction is the solution of the inverse problem of the above image formation processes. Given the spatiotemporal sparsity of the dynamic scene, which holds in many if not most natural scenes, a CS-based reconstruction algorithm can successfully decode the spatiotemporal mixing in the vertical axis of the streak camera and retrieve spatiotemporal information. Within a single exposure, CUP can capture up to 350 consecutive temporal frames at 100 billion frames per second.

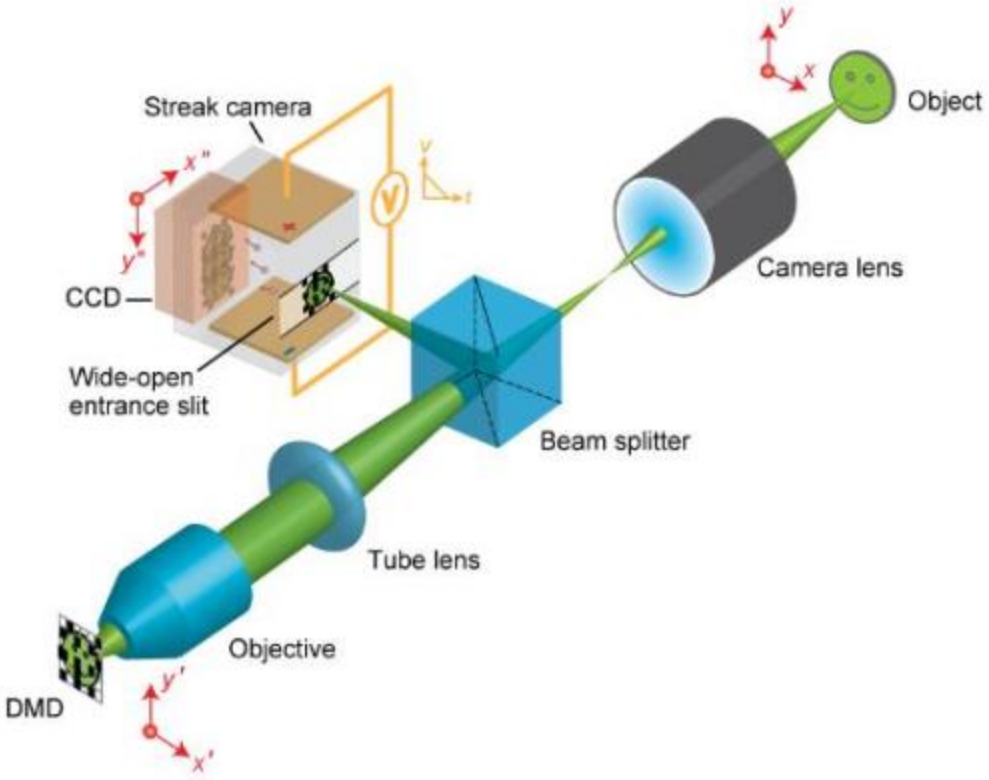

Figure 1.6 Optical schematics of CUP [12].

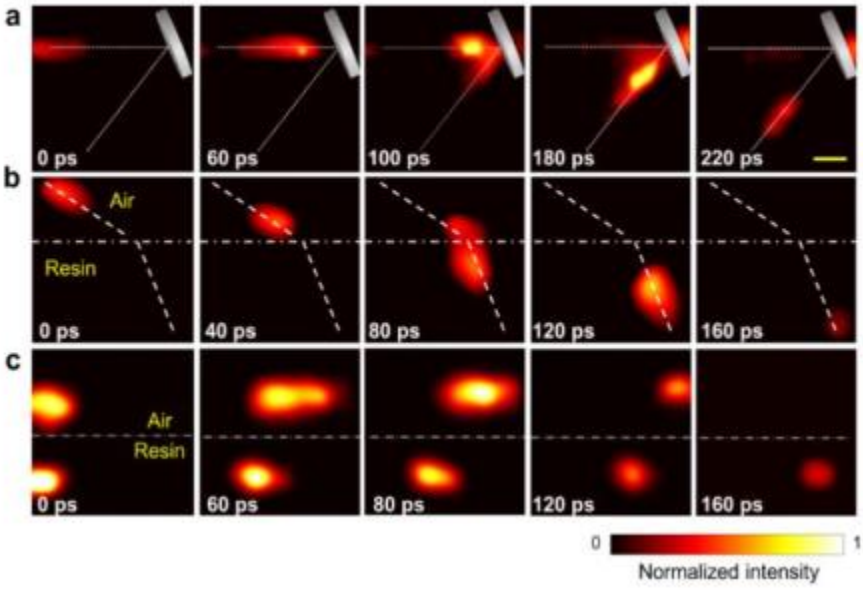

Figure 1.7 CUP of light propagation [12]. (a) Laser pulse reflected from a mirror. **(b)** Laser pulse refracted from an air-resin interface. **(c)** Laser pulses racing in air and resin. Scale bar, 10 mm.

Using CUP, many transient light-speed phenomena have been visualized, including the propagation, reflection, and refraction of a short laser pulse in space, faster-than-light propagation of non-information, and color-resolved fluorescent excitation and emission and dynamic volumetric imaging. In CUP results presented in **Figure 1.7**, reconstructed light reflection and refraction obey Snell's law (**Figure 1.7 a**, **Figure 1.7 b**). For the first time, this technique enables video recording of photon propagation at a temporal resolution down to tens of picoseconds. Moreover, light racing in two different media (air and resin) is imaged and two light speeds are quantitatively compared in air and resin by tracing the centroids of laser pulses in the time lapse video (**Figure 1.7 c**). The calculated results, light velocity in air = (3.1 ± 0.5) × 108 m/s and light velocity in resin = (2.0 ± 0.2) × 108 m/s, are very close to the theoretical values.

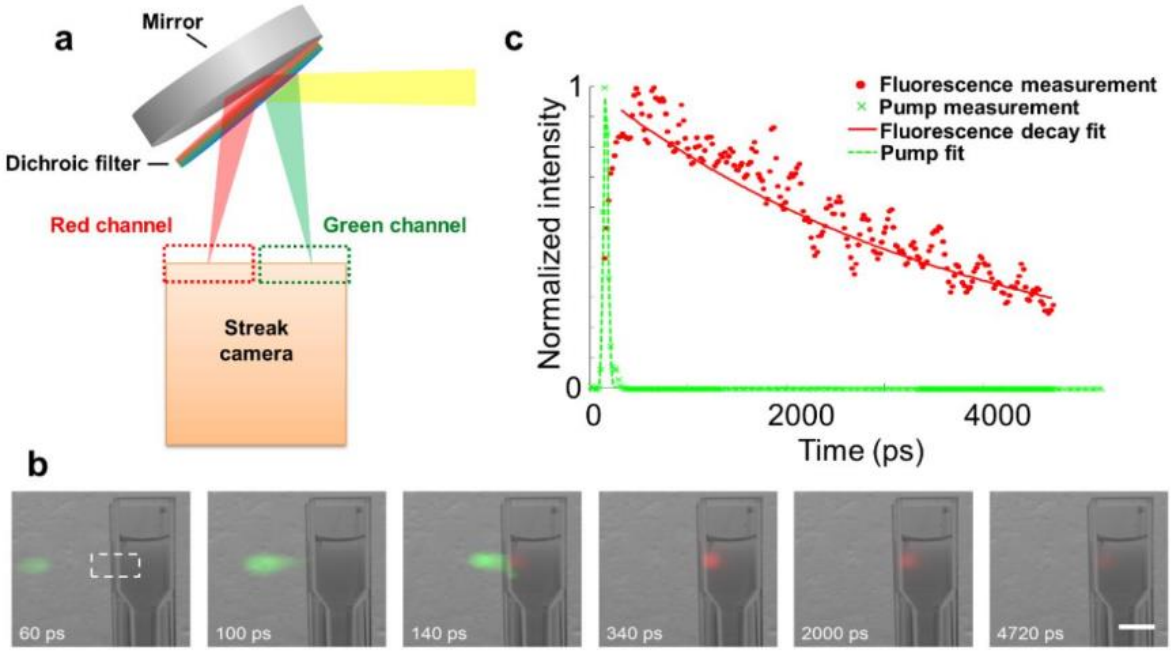

Figure 1.8 Multicolor CUP [12]. (a) Custom-built spectral separation unit. **(b)** Representative temporal frames of a pulsed laser-pumped fluorescence emission process. The pulsed pump laser and fluorescence emission are pseudo-colored based on their peak emission wavelengths. To explicitly indicate the spatiotemporal pattern of this event, the CUP-reconstructed frames are overlaid with a static background image captured by a monochromatic CCD camera. **(c)** Time-lapse pump laser and fluorescence emission intensities averaged within the dashed box in b. The temporal responses of pump laser excitation and fluorescence decay are fitted to a Gaussian function and an exponential function, respectively. Scale bar, 10 mm.

To enable multicolor CUP, a spectral separation module in front of the streak camera can be added. As shown in **Figure 1.8 a**, a dichroic filter (562 nm cut-on wavelength) is mounted on a mirror at a small tilt angle (~5°). The light reflected from this module is divided into two beams according to the wavelength: green light (wavelength < 562 nm) is directly reflected from the dichroic filter, while red light (wavelength > 562 nm) passes through the dichroic filter and bounces from the mirror. Using the multicolor CUP system, a pulsed-laser-pumped fluorescence emission process is imaged. A fluorophore, Rhodamine 6G, in water solution was excited by a single 7 ps laser pulse at 513 nm. Some representative temporal frames are shown in **Figure 1.8 b**. In addition, the time-lapse mean signal intensities within the dashed box in **Figure 1.8 b** for both the green and red channels (**Figure 1.8 c**) are calculated. Based on the measured fluorescence decay, the fluorescence lifetime was found to be 3.8 ns, closely matching a previously reported value.

Currently, CUP relies on the unconstrained two-step iterative shrinkage/thresholding (TwIST) algorithm to reconstruct the event data cube. CUP is a two-dimensional ultrafast dynamic passive image platform and has three characteristics. First, it features high imaging speed and can reach picosecond time resolution. Second, it is a single-shot imager and does not scan the sample, thereby minimizing the exposure to the sample, which is beneficial to live tissue imaging. Third, it is passive imaging and does not require active illumination. As a consequence, it is suitable for live cell imaging.

2. Motivation

Space and time are two key physical dimensions in modern metrology. In bio-imaging, significant advances have been made to improve the spatial resolution in microscopy as recognized by the 2014 Nobel Prize in chemistry. The nanoscopy tools can break the diffraction limit of light. High-temporal-resolution imaging is equally important and has become possible because of the ultrafast imagers such as CUP. Ultrafast imagers operate at an ultrafast speed to resolve fast dynamics and provide quantitative measurements to the fast dynamics at molecular levels in order to explore the underlying mechanisms which occur at various timescales ranging from microseconds, such as ligand binding and protein folding, to femtoseconds, such as molecular vibration and energy transfer. Ultrafast imaging is essential for observation and characterization of fast dynamic events.

Despite high sensitivity, the common drawback of current time-domain and frequency-domain FLIM imagers is their dependence on repetitive measurements. For example, to acquire a 2D image, a time resolved confocal microscope or streak-camera has to scan pointwise and linewise, respectively. To avoid motion artifacts, the sample must remain static during data acquisition. Widefield FLIM systems acquire spatial data in parallel. However, to achieve high temporal resolution, they still need to temporally scan either a gated window or detection phase, or they have to use TCSPC which requires a large number of repetitive measurements to construct a temporal histogram. Limited by this bottleneck, current FLIM systems acquire lifetime images at only a few frames per second and this slow frame rate restricts these systems from imaging transient biological events, such as neural spiking, calcium oscillation and fast protein folding which typically occur within a few milliseconds. For example, most current endoscopic FLIM systems suffer from either motion artifacts because of prolonged acquisition time or poor lifetime estimation accuracy owing to insufficient temporal sampling. Similarly, to measure the intracellular temperature distribution in live cells, despite high sensitivity and functional independency, current FLIM systems fail to provide enough temporal resolution to image transient cellular thermogenesis, such as temperature fluctuation under drug stimulus. As is illustrated in **Figure 2.1**, high temporal resolution has attracted attention recently because of the fast dynamics in the life sciences. The fast dynamics events require up to MHz and GHz imaging speed to resolve those events. **Figure 2.2** illustrates that in order to resolve fast biodynamics such as molecular vibration, protein folding and neuron spiking, ultrafast imagers such as CUP are required. Current FLIM systems cannot resolve the details due to the repetitive imaging and mechanical scanning.

Moreover, most of the fast dynamics are essentially scholastic processes and require snapshot imaging. However, current FLIM systems are based on repetitive excitation and detection cycles, and they assume that the target fast dynamics are strictly repeatable in each excitation cycle such as the fluorescence decay process. Consequently, they cannot be applied to detect random processes such as neuron spiking which are beyond prediction and protein folding processes which can be easily affected by the environments.

To increase the frame rate of FLIM, a number of approaches have been developed in the past decade. A common strategy is to reduce the number of necessary measurements. For example, rather than measuring the entire fluorescence decay with dense temporal samplings, a widefield FLIM imager can measure signals in only a few temporal channels when using a gated imager intensifier [13–17] or measure only a limited number of photons when using TCSPC [18]. However, it compromises the fitting accuracy in lifetime estimation. Similarly, in the spatial domain, instead of scanning the entire field of view, a confocal FLIM system can adopt a random-access strategy and scan only regions of interest [19, 20]. Nonetheless, this method needs a complete pre-scan to acquire the prior knowledge about fluorophores' distribution. More problematically, the objects must stay within the boundary of this pre-defined area during the entire data acquisition. Besides the aforementioned strategy, there are a few other techniques that also enable fast FLIM. Won et al. replaced the TCSPC process with analog mean delay in a confocal FLIM system and increased the imaging speed by two orders of magnitude [21]. However, this method works only for fluorescence lifetimes with a single exponential decay. Additionally, due to the reliance on spatial scanning, the frame rate is limited to 7.7 fps (image size, 158×127 pixels). In another representative work, Tsikouras et al. employed a 2D-to-1D fiber bundle to rearrange a 2D input image into a 1D output line, followed by passing the resultant signals to the entrance slit of a streak camera [22]. Despite snapshot acquisition format, the spatial resolution is low—the lifetimes can be simultaneously measured in only 10×10 pixels.

Therefore, an innovative ultrafast imaging method is required to integrate FLIM into ultrafast bio-dynamics detection. This master project explores the approach, dubbed compressed FLIM, to overcome these limitations by incorporating the paradigm of compressed sensing into FLIM. The resultant system will enable widefield, depth-resolved lifetime imaging at up to 100 fps, approximately two orders of magnitude faster than that allowed by current FLIM imagers.

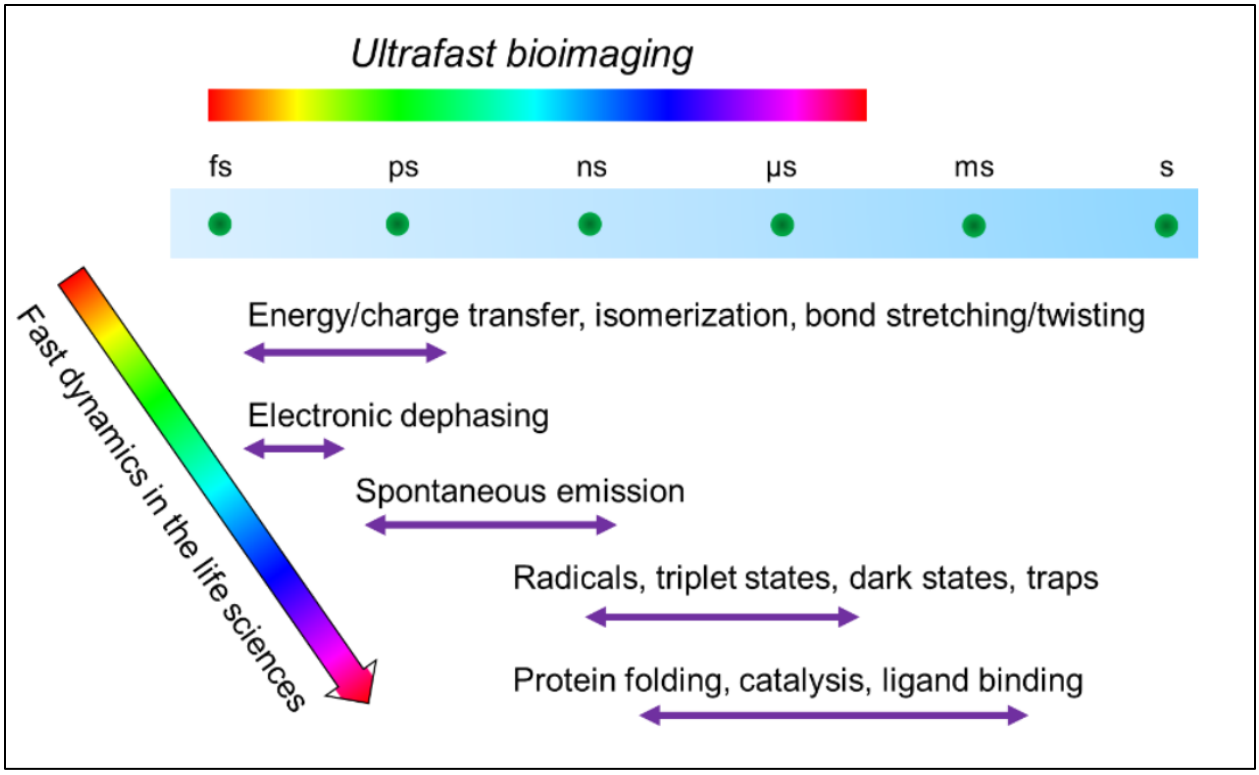

Figure 2.1 Fast dynamics in life sciences. Ultrafast bioprocesses occur at a varying timescale ranging from microseconds, such as ligand binding, to femtoseconds, such as molecular vibration. Resolving these fast dynamics requires ultrafast bioimaging.

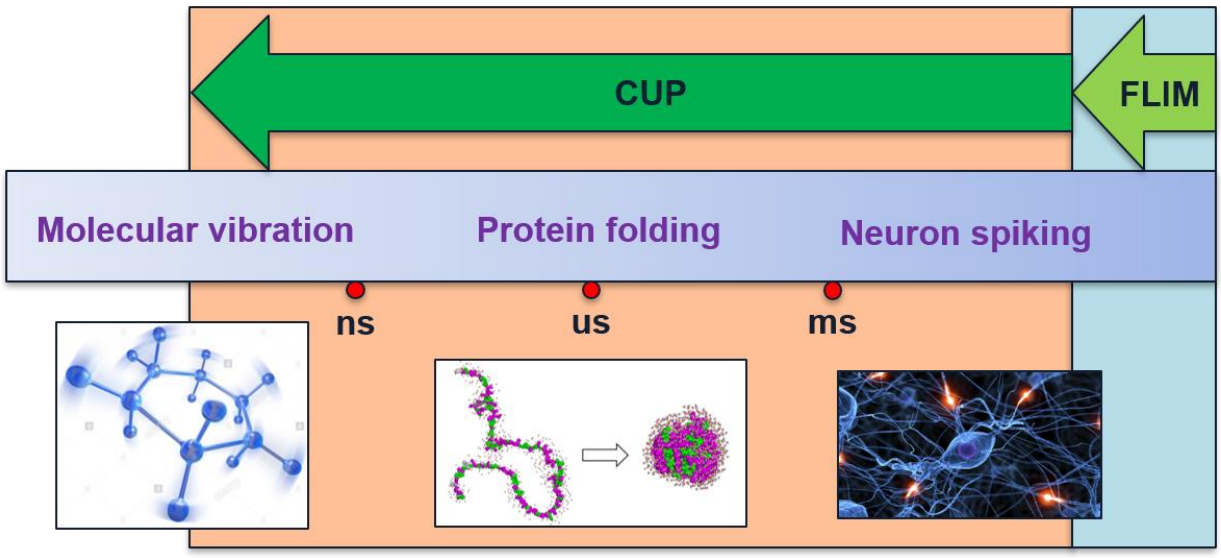

Figure 2.2 Ultrafast biodynamics requires ultrafast imager such as CUP. Current FLIM cannot resolve those fast dynamics.

3. Rationale

3.1. Overview of compressed FLIM

To increase the FLIM imaging speed and capture the ultrafast bio-dynamics, FLIM-CUP, dubbed compressed FLIM, is built upon compressed ultrafast photography (CUP) technology. However, rather than targeting photographic applications, compressed FLIM is tailored for high-resolution fluorescence microscopy. Compressed FLIM features two creative technical aspects. First, it significantly increases the resolution of recovered lifetime images by employing a dual-camera setup detection which enables high-resolution lifetime imaging. In the initial work of CUP, the spatial resolution of CUP degrades from the diffraction limit. The loss of resolving power therefore hinders the application of CUP in fluorescence microscopy. To recover high-resolution lifetime images, compressed FLIM will employ a dual-camera setup—besides the original streak camera, a reference intensity camera is added to capture a time-integrated image, which then will be used as the spatial constraint during image reconstruction. By leveraging the rich information contained in the temporally sheared and integrated images, the quality of reconstructed lifetime maps can be significantly improved. Second, it enables high-speed lifetime imaging by adopting a scientific CMOS as the streak camera's photon detector and optimizing its readout. Capable of depth resolved lifetime imaging at 100 fps, the resultant system is approximately two orders of magnitude faster than state-of-the-art FLIM imagers; last but not least, it alleviates the photobleaching at the sample. Compared with flood excitation in conventional confocal and widefield FLIM imagers, the average illumination power at the sample is reduced to a large degree in compressed FLIM due to the single shot for one given field of view, decreasing the photobleaching rate and thereby permitting a prolonged observation time window.

Specifically, compressed CUP uses CUP as the detector in FLIM and combines CUP and FLIM into one system. It reconstructs a discrete x-y-t (x, y, spatial coordinates; t, time) estimate of the target fluorescence decay data cube from a single 2D streak image and one reference time-integrated image. The reconstruction is accomplished using optimization algorithms such as fast iterative shrinkage/thresholding algorithm (FISTA) and two-step iterative shrinkage/thresholding (TwIST) algorithm. It can detect the fluorescence decay of multiple fluorophores and by single-shot detection; the mixed fluorescence delay dynamic event is then estimated though compressed sensing and finally through post-processing the lifemap is generated.

As a single-shot fluorescence lifemap imager, compressed FLIM overcomes speed limitation and will lead to a new generation of widefield lifetime imagers. It will enable high-speed recording of fluorescence lifetimes in transient biological events. Since compressed FLIM measures the entire fluorescence decay rather than a few temporal channels, it will allow decomposition of multiple fluorophores through multi-exponential curve fitting.

Compressed FLIM consists of two subsystems: (1) the fluorescence microscope, (2) the CUP camera. As a computational imaging approach, compressed FLIM operates in two steps: data acquisition and image reconstruction. In data acquisition, first, the sample is excited via the fluorescence microscope. Then, the CUP camera measures the fluorescence. In image reconstruction, a graphics processing unit (GPU) assisted accelerated compressed sensing reconstruction algorithm, developed from the two-step iterative shrinkage/thresholding (TwIST) algorithm [23], processes the images and performs lifetime measurements in real time.

In data acquisition, the CUP camera in compressed FLIM employs dual-camera detection and acquires three different views of the sample's fluorescence. One view, akin to the fluorescence intensity microscope, captures a spatially unencoded and temporally unsheared fluorescence intensity image temporally integrated over the exposure time. In contrast, the other two views capture two spatially encoded and temporally sheared fluorescence images under the compressed sensing paradigm. Upon excitation via an epi-fluorescence microscope, a beam splitter divides the fluorescence from the microscope side port into two components. The transmitted component is imaged by an external reference camera to form the temporally unsheared view. The reflected component is relayed to a digital micromirror device (DMD) through a 4f imaging system composed of a tube lens, a mirror, and a stereoscope objective. A static, pseudorandom, binary pattern is displayed on the DMD to encode the image. Each encoding pixel is turned either on (tilted –12° with respect to the DMD surface norm) or off (tilted 12° with respect to the DMD surface norm) and reflects the incident fluorescence in one of the two directions. Two reflected fluorescence beams, masked with complementary patterns, are both collected by the same high numerical aperture stereoscope objective lens and enter corresponding sub-pupils at the objective's back focal plane. The fluorescence from these two sub-pupils is then imaged by two tube lenses, folded by right-angle prism mirrors to form two complimentary channel images in separate horizontal areas at the entrance port of a streak camera. To accept the encoded 2D image, this entrance port is fully opened (~5 mm width) to its maximum, exposing the entire photocathode to

the incident light. The encoded fluorescence signals are temporally sheared along the vertical y axis according to the time of arrival inside the streak camera and then land at different spatial positions along the vertical axis to form an image at the streak camera's output port. The image is then measured by a 2D detector array—an internal CMOS camera. The photons are temporally integrated within the camera's exposure time and spatially integrated within the camera's pixel, forming two temporally sheared views. The dual-channel detection setup features a 100% light throughput, saving every photon in low-light conditions.

In image reconstruction, compressed FLIM first recovers the dynamic fluorescence decay scene from the three-view data by exploiting the spatiotemporal sparsity of the event in most of the experimental conditions, then performs the exponential lifetime fitting at each spatial point and produces the reconstructed lifetime image. As in the dual-camera detection, the temporally unsheared view recorded by the external reference camera functions as the spatial constraint in the image reconstruction and improves the reconstructed image spatial resolution. In addition, two complementary temporally sheared views are formed from both on and off pixels of the DMD and this design prevents any loss of information from spatial encoding, which enriches the observation and is advantageous for compressed sensing-based reconstruction. Thus, the dual-camera detection, the dual complementary masking, and the three-view joint reconstruction in the compressed FLIM image reconstruction strategy yield high-resolution lifetime imaging.

The external reference camera and streak camera are synchronized to perform high-speed lifetime imaging. Upon every single excitation, one time-integrated fluorescence intensity image without shearing and encoding is acquired from the external reference camera, and one spatially encoded and temporally sheared image recording the two views is acquired from the streak camera. The reconstructed lifetime image is produced in real-time from the previous two acquired images. The frame rate of compressed FLIM depends on the image acquisition speed of both the external reference camera and streak camera and can reach 100 fps. The spatial resolution, determined by the numerical aperture of the objective lens, will be in a submicron range, providing a resolving power to uncover the folding heterogeneities inside a cell.

Compressed FLIM's unique paradigm of data acquisition and image reconstruction brings several prominent advantages. First, facilitated by the streak camera, the compressed FLIM system can image a nonrepetitive dynamic scene at 100 fps depth-resolved lifetime imaging with a single-snapshot measurement, circumventing the necessity of repetitive measurements by the pump-

probe technique [24–26]. Second, compressed FLIM does not need the specialized active illumination required by other single-shot ultrafast imagers [27–29], enabling passive imaging of dynamic scenes. Third, compared with other streak camera-based single-shot ultrafast imaging methods [30, 31], the compressed FLIM system has a light throughput of nominally 100% (excluding imperfect losses from the optical elements).

3.2. Operating principle of compressed FLIM

Compressed FLIM involves physical data acquisition and computational image reconstruction. As discussed previously [12], the system forward model can be expressed as

$$E = \boldsymbol{O}I, \tag{3.1}$$

where $\boldsymbol{O}$ represents the system forward response of the streak camera, E is the streak camera measurement, and I is the target fluorescence decay data cube.

To impose a spatial constraint, we adopt the dual camera design [32] and split the fluorescence toward two channels. One channel is for the streak camera to conduct the compressed sensing of the 3D decay data cube. Fluorescence in the other channel is directly captured by the high-resolution reference intensity camera to operate the uniform sampling along the temporal dimension, which can be expressed as

$$\tilde{E} = \widetilde{\boldsymbol{O}}I, \tag{3.2}$$

where $\widetilde{\boldsymbol{O}}$ denotes the system forward response of the reference intensity camera.

With this spatial constraint, we obtain two measurements which obey the observation model as follows:

$$E' = \boldsymbol{O}'I + \boldsymbol{v}', \tag{3.3}$$

where $E' = \begin{bmatrix} E \\ \tilde{E} \end{bmatrix}, O' = \begin{bmatrix} O \\ \tilde{O} \end{bmatrix}, v' = \begin{bmatrix} v \\ \tilde{v} \end{bmatrix}$, and v and $\tilde{v}$ are the additive noise of the two detectors.

We recover the data cube by solving the following inverse problem:

$$\hat{I} = \arg\min_{I} \|E' - \boldsymbol{O}'I\|_2^2 + \beta\Phi(I), \tag{3.4}$$

where $\Phi(I)$ is a regularization term imposing sparsity and β is the weighing factor between the fidelity and sparsity.

To further impose space and intensity constraint, we construct the new constrained solver

$$\hat{I}_s = \arg\min_{I \in M, I > s} \|E' - \boldsymbol{O}'I\|_2^2 + \beta\Phi(I). \tag{3.5}$$

Here, *M* is a set of possible solutions confined by a spatial mask extracted from the reference intensity image and defines the zone of action in the reconstruction. This spatial constraint improves the image resolution and accelerates the reconstruction. *S* is the low intensity threshold constraint to reduce the low-intensity artifacts in the data cube. This space and intensity constrained solver can be coupled with various optimization algorithms such as DeSCI [33] and GAP-TV [34]. After the reconstruction, least square exponential fitting is applied to data cube along the temporal dimension at each spatial point to extract the lifemap image. **Figure 3.1** illustrates the system's operating flowcharts, i.e., the forward model of compressed FLIM.

To accelerate the reconstruction, we implement the algorithm with CUDA parallel programming framework on two NVIDIA Tesla K40 GPUs and we also leverage the computing power of the Illinois Campus Cluster. This acceleration reduces typical reconstruction time from tens of minutes to seconds.

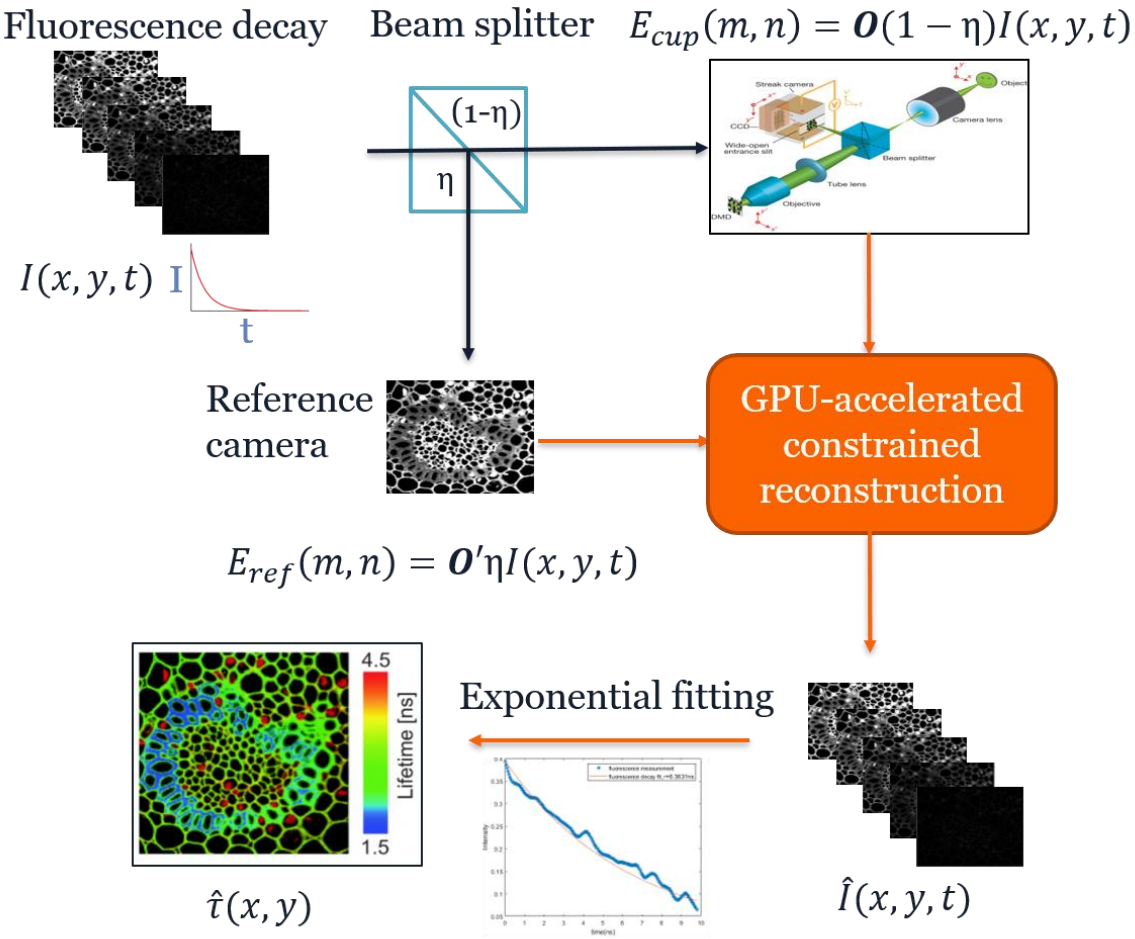

Figure 3.1 Compressed FLIM operating flowcharts for the forward model of compressed FLIM. $\boldsymbol{I}$ is the fluorescent decay data cube. η is the transmission ratio of beam splitter. $\boldsymbol{O}'$ refers to the forward response of reference camera. $\boldsymbol{O}$ represents the forward response of CUP camera. $\hat{\boldsymbol{\tau}}$ is the estimated lifemap reconstructed by the compressed FLIM. $\boldsymbol{E}$ represents the optical energy. $\boldsymbol{E_{cup}}$ is the optical energy distribution captured by the compressed FLIM camera. $\boldsymbol{E_{ref}}$ represents the optical energy distribution captured from the reference camera. $\boldsymbol{m}$ and $\boldsymbol{n}$ are the pixel index. The dynamic scene is the fluorescence decay from a cell slide upon excitation.

4. Results

4.1. Compressed FLIM system details

In the compressed FLIM system, we use an epi-fluorescence microscope (Olympus IX83) with a 60× 1.42 NA oil immersion objective (PLAPON 60XO, Olympus) for H&E slides, stained neurons and fluorescent beads. We use a 4× 0.1 NA objective (PLN 4X, Olympus) for FRET imaging in vitro. We use 10× 0.3 NA objective (UPLFLN 10X2, Olympus) for fluorescent tissue paper imaging. We use 513 nm picosecond pulse laser (Genki-XPC, NKT Photonics) as excitation source, a 532 nm dichroic mirror (ZT532rdc, Chroma) and a 590/50 nm band-pass emission filter (ET590/50m, Chroma). We use an engineered diffuser (ED1-C20-MD, Thorlabs) to get uniform excitation and guide the illumination through an L-shape-fluorescence illuminator (IX3-RFAL) to the sample stage. Upon excitation, the fluorescence from the sample is guided out of the microscope via left side-port. Following the intermediate imaging plane, a beam splitter (BSX16, Thorlabs) transmits 10 percent of the light to an external reference camera (CS2100M-USB, Thorlabs). The rest of the reflected light is imaged to a DMD (DLP LightCrafter 6500, Texas Instruments) through a 4f system consisting of a tube lens (AC508-100-A, Thorlabs) and a stereoscope objective (MV PLAPO 2XC, Olympus; numerical aperture, 0.50). The pattern displayed in DMD is 3×3 binning. The spatially encoded images are relayed to the entrance port of a streak camera (C13410-01A, Hamamatsu) through two 4f systems containing the same stereoscope objective, tube lenses (AC254-75-A, Thorlabs), and the right-angle prism mirrors (MRA10-P01, Thorlabs; #47004, Edmund Optics). The shearing velocity of the streak camera is set to v = 0.132 mm/ns (time window = 100 ns) for H&E slides, stained neurons, fluorescent beads, FRET imaging in vitro and fluorescent tissue paper imaging. The spatially encoded, temporally sheared images are acquired by an internal CMOS camera (ORCA-Flash4.0 V3, C13440-20CU, Hamamatsu) with a sensor size of 1344 (H) × 1016 (V) pixels (1×1 binning; pixel size d = 6.5 μm). Compressed FLIM is synchronized with a digital delay generator (DG645, Stanford Research Systems). In practice, the reconstructed data cube size, $N_x \times N_y \times N_t$, was limited by the size of the internal CMOS camera, $N_R \times N_C$, where N_R and N_C are the number of rows and columns. The reconstructed datacube size $N_x \times N_y \times N_t$ is limited by the pixel format $N_R \times N_C$ of the internal CMOS camera where N_R and N_C are the number of rows and columns. In compressed FLIM, the two complementary views occupy two horizontally equal areas of the internal CMOS camera simultaneously. Consequently, N_x must satisfy $N_x \leq N_C/2$. In addition, because the temporal

shearing occurs vertically, N_y and N_t must satisfy $N_y + N_t - 1 \leq N_R$. With a fully opened entrance port, each temporal frame has an approximate size of $N_x \times N_y = 450 \times 500$, which provides an approximate sequence depth of $N_t = 600$. Thus, the typical reconstructed datacube size in compressed FLIM is $N_x \times N_y \times N_t = 450 \times 500 \times 600$.

Figure 4.1 illustrates the compressed FIIM system configuration. **Figure 4.1 a** presents the excitation part using the epi-fluorescence microscope (Olympus, IX83). A beam splitter divides the total fluorescence into two parts. One part goes to the external reference camera. The second part is collected by the CUP camera. **Figure 4.1 b** presents the top view of the CUP camera in the compressed FLIM system. After encoding, shearing and spatiotemporal integration, the final spatially encoded, temporally sheared streak image is detected by the internal CMOS camera inside the streak camera.

4.2. Fluorescent decay of a tissue paper

To validate the extremely high temporal and high spatial resolution imaging capability of the compressed FLIM, we image a fluorescent tissue paper. The tissue paper is marked by a yellow highlighter. We illuminate the tissue paper with 513 nm picosecond laser purchased from NKT Photonics. The pulse width is 7 ps and is negligible compared with the lifetime of the highlighter fluorophore (around 10 ns). Time window is 100 ns based on the lifetime of the highlighter fluorophore to optimize the signal-to-noise ratio of the streak image of the fluorescence signal. Under single shot, the fluorescence decay data cube is recovered using the TwIST algorithm. Exponential fitting is performed at each spatial point along the time axis to extract the lifemap of the target fluorophore. **Figure 4.2** illustrates the recovered exponential decay process. The time resolution is tens of picoseconds. Spatial resolution is 500 × 450 pixels. **Figure 4.3** shows the recovered fluorescence decay process at one spatial point in the tissue paper marked in the 0 ns frame. Single exponential fitting is performed. The fitting matches well with the recovered decay trace by compressed FLIM.

4.3. Hematoxylin and Eosin (H&E) Slide

4.3.1.H&E slide

To demonstrate the ability of compressed FLIM to perform high-resolution lifetime imaging, we use it to image the hematoxylin and eosin (H&E) stained clinical histopathology slide. The specimen must be stained in order to make the cells visible and microscopically observable in a processed, sectioned and mounted tissue slide. Histology slide staining is to stain the slide with

two or more contrasting dyes to distinguish different entities in the tissue slides. Dyes prove to have an affinity for structures of cells and tissues because of the attraction forces between the dye molecules and cell and tissue molecules. These kinds of affinity depend on dye structures, shape of the dye molecules, charge distributions and solvent characteristics and generally are stronger than those of the solvent molecules. Charge distribution determines the force characteristics of the stains. Acidophilic stains have the acidic pH (<7) and therefore cell and tissue molecules will bind H^+ and present positive. By contrast, basophilic stains have the basic pH (>7) and tissue molecules will give up H^+ and present negative. H&E stain is the most widely used stain in histology. Eosin will tend to stain connective tissues and extracellular materials and hematoxylin will stain cell nuclei. Eosin is a red, acidic and negative stain which binds to acidophilic substances such as positively charged amino-acid side chains (e.g. lysine, arginine). It stains the cytoplasmic materials such as connective tissue and collagen. Hematoxylin is a purple stain which marks the nuclear material in the nucleus. It binds to aluminum so has a positive charge distribution and hematoxylin/aluminum binds negative molecules in cells and tissues such as chromatic material in nucleus. As the standard stain in histology, it makes the cell nucleus visible and presents their activity states. These two stains present a sharp contrast to distinguish various parts in tissues. Malignancy of the clinical sample mostly depends on H&E slide for diagnosis. Progressive and regressive staining systems are two common staining procedures of H&E staining. In progressive staining, staining time is controlled to determine the intensity of hematoxylin, and hematoxylin to tissue is point-to-point staining. In regressive staining, hematoxylin overstains the whole tissue intentionally and background is removed later by means of differentiation which decreases the binding ability of hematoxylin to the tissue point sites. Specific staining refers to the staining of the nuclei and DNA. Non-specific staining, also called background staining, refers to cytoskeleton structures. The commercially available hematoxylins are Harris, Mayers, Gill I, II, III and others.

4.3.2. Compressed FLIM of H&E slide

We image the H&E slide using 513 nm picosecond laser under single shot and the reconstruct the fluorescent event data cube and perform the exponential fitting to extract the lifetime map under the given field of view (FOV). In the reconstructed lifetime image **Figure 4.4 b**, the dark circles represent the nuclei because the hematoxylin does not fluoresce under 513 nm excitation and bright

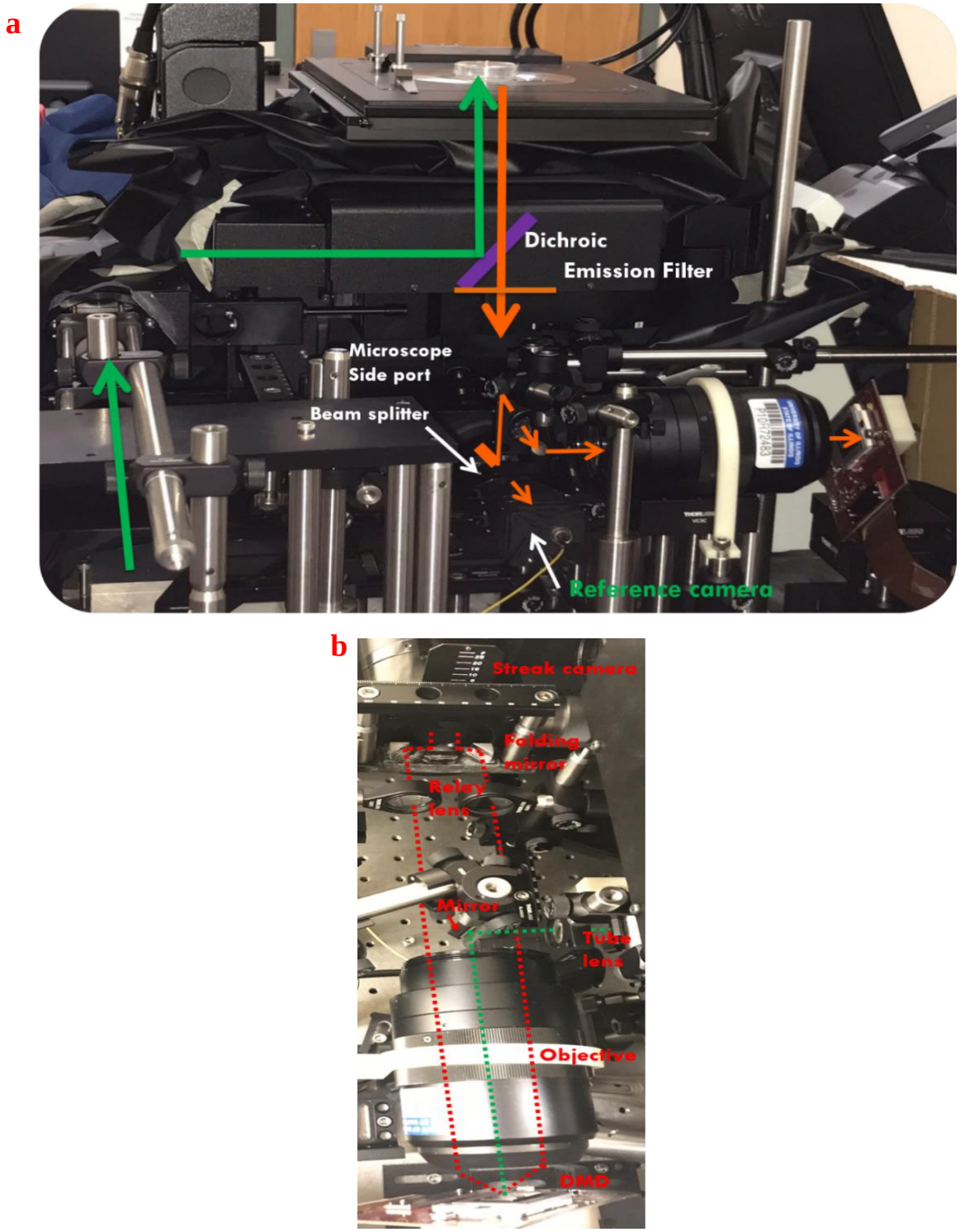

Figure 4.1 Compressed FIIM system configuration. (a) Epi-fluorescence microscope structure. **(b)** Top view of compressed FLIM detection structure.

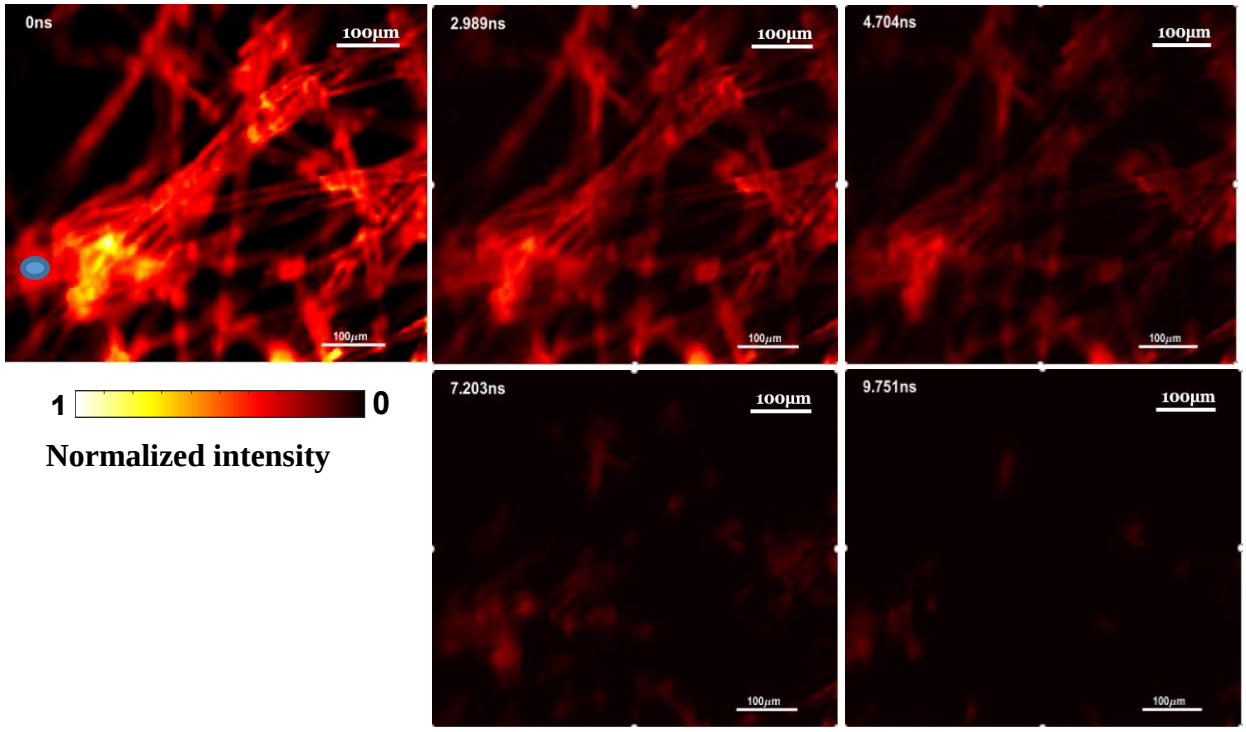

Figure 4.2 Recovered fluorescence decay process of fluorescent tissue paper by compressed FLIM. Five frames at five different time stamps are shown above.

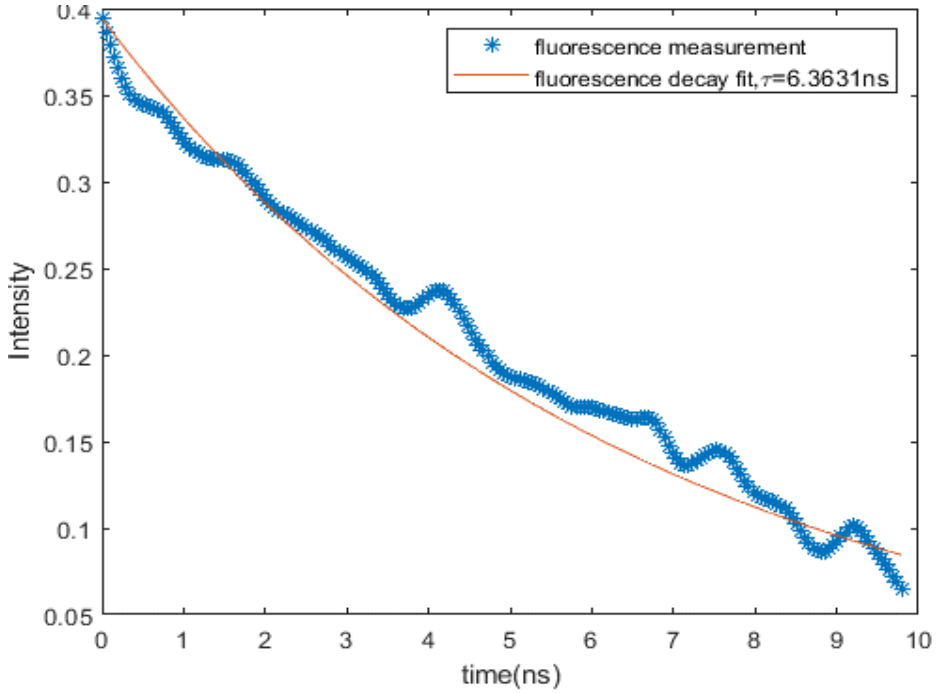

Figure 4.3 Recovered fluorescence decay process at the blue spatial point marked in the fluorescent tissue paper by compressed FLIM. Five frames at five different time stamps are shown above.

areas with varying fluorescence lifetime are based on quenching of eosin. **Figure 4.4 a** is the intensity image captured by the reference camera. To further validate the result, we image the H&E slide by conventionally line-scanning streak camera fluorescence lifetime imaging and get the ground truth lifetime image. The reconstructed lifetime image matches well with the ground truth image.

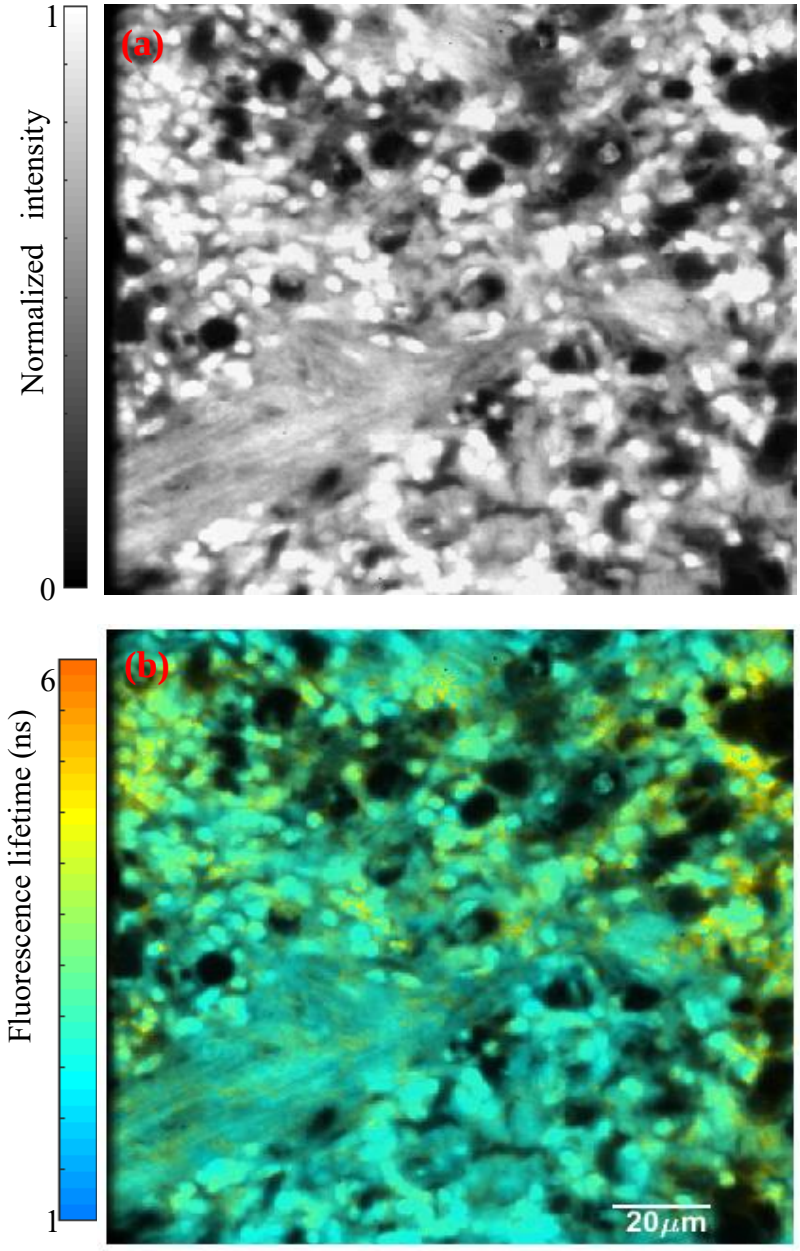

Figure 4.4 Lifetime imaging of compressed FLIM. (a) Intensity image of hematoxylin & eosin (H&E) clinical histopathology slides. **(b)** Reconstructed lifetime image of hematoxylin & eosin (H&E) clinical histopathology slides.

4.4. Cytoskeleton structure labelled with two fluorophores

4.4.1. Inverse problem: fluorophores temporal unmixing

Fluorescence labeling is the standard way to differentiate intracellular structures. Multiple fluorophores are used to label different subcellular structures such as nuclei and mitochondria. Consequently, the ability to image them simultaneously (multiplexing) has been critical for understanding a variety of biological process.

Generally, two methods have been applied to achieve multiplexing. The first is the spectrum-based method, which is a conventional and widely adopted method and separates fluorophores with their emission spectra. Fluorophores with different emission wavelengths are selected and the overall combined fluorescence signals are then filtered into multiple channels and imaged separately, which are then superimposed to obtain the final image. This method can differentiate fluorophores accurately but is limited by the fluorophore selections, and the elongated imaging process might cause phototoxicity to the sample, which does harm to live cell imaging. Another spectrum-based method postprocesses the mixed spectral measurement to unmix the combined signals with algorithms such as least squares algorithms. It is restricted by the limited bandwidth of the filter and it is difficult to resolve more than two fluorophores with highly overlapping spectra without sophisticated unmixing algorithms. The second method is based on the fluorescence lifetime, an identifying characteristic separate from the emission spectrum that can be used to distinguish fluorophores. It overcomes the aforementioned limit in spectrum-based method by using the additional fluorescence lifetime dimension. This unlocks the potential of separating multiple fluorescent probes. This approach relies on the fact that fluorophores exhibit a time-resolved decay pattern that can be easily characterized and is very often mono-exponential with a decay rate of $1/\tau$. Lifetime multiplexing assumes that each fluorophore present has a characteristic lifetime decay. A mixture of n fluorophores, when excited by an impulse, will exhibit time-dependent fluorescence emission $\boldsymbol{f(t)}$ that is a sum of the respective fluorophore decays:

$$f(t) = a_0 + a_1 \exp\left(-\frac{t}{\tau_1}\right) + a_2 \exp\left(-\frac{t}{\tau_2}\right) \ldots + a_n \exp\left(-\frac{t}{\tau_n}\right). \quad (4.1)$$

Although sophisticated numerical methods exist to obtain the amplitudes and lifetimes from a mixture of fluorophores, simplification is achieved in many cases when the lifetimes can be determined experimentally in advance. Assuming known lifetimes (τ_1, …, τ_n), amplitude

components (a_1, …, a_n) can be determined by directly fitting ***f*(*t*)**. It provides a simple way to unmix or separate the signal from multiple fluorophores and has been used for fluorescence microscope and immunofluorescence imaging. Lifetime contrast is independent of fluorescence intensity and emission spectrum and is more robust than spectrum-based methods which are highly influenced by the fluorescence intensity. In recent studies, Leica Microsystems has successfully differentiated subcellular proteins Vimentin and Tubulin labeled by highly overlapped fluorophore pairs Alexa Fluorophore 555 and Alexa Fluorophore 546 respectively. Raymond et al. did the in vivo lifetime multiplexing of highly overlapped near infrared fluorophore pairs Osteosense and Kodak X-Sight in nude mice [35].

In spectral imaging, a spectral imager captures both the spatial and spectral information of the scene, providing a 3D (x, y, λ) (x, y, spatial coordinates; λ, wavelengths) data cube for multivariable analysis. Using the spectral information, we can separate different fluorophores based on their spectral signatures. So spectral unmixing solves the inverse problem of spectral imaging. Similar to spectral unmixing, in lifetime-based temporal unmixing, compressed FLIM measures the 3D (x, y, t) (x, y, spatial coordinates; t, time) data cube. In spectral unmixing, spectral component matrix is given as the prior information which derives from the emission spectra of the mixed fluorophores components. Likewise, decay component matrix in temporal unmixing results from the fluorescence lifetime of each component. Then in discrete form we construct the inverse problem,

$$y = \mathbf{A}x + \mathbf{W}, \tag{4.2}$$

where y denotes the measurement from compressed CUP system, **A** is the prior decay matrix of mixed fluorophores, x is the fluorophore concentration matrix and **W** is the noise matrix. The time-dependent measurement y is linearly related to the amplitudes of the individual decay components decay where the columns of **A** are the normalized, time-dependent mono-exponentially intensity decays and x are the component amplitudes.

Assuming known lifetimes (τ_1, …, τ_n) and each fluorophore present has a characteristic lifetime mono-exponential decay, the inverse problem then evolves into a least squares problem with constraints:

$$\text{Minimize } \|\boldsymbol{A}x - y\| \text{ subject to } x \in \mathcal{C}, \tag{4.3}$$

where *C* is a convex set as prior knowledge of the vector x to be estimated.

Several convex sets perform as the regularization to rule out certain unacceptable approximation of x. They arise as prior knowledge of the vector x to be estimated and are the same as determining the projection of y on a set more complicated than a subspace. Typically, they form a convex optimization problem with no analytical solution such as non-negativity constraints on variables, variable bounds, probability distribution and norm ball constraint (trust region).

Given x as the amplitude matrix in temporal unmixing problem, several kinds of constraints can be applied to the least squares. First, add the 2-norm penalty to the objective function and form the solvent ℓ_2-regularized least squares problems:

$$\min_x \ \|\boldsymbol{A}x - y\|_2^2 \ + \ \lambda\|x\|_2^2 , \tag{4.4}$$

where λ is the regularization parameter. 2-norm penalty seeks for an approximate solution of **A**x ≈ y with small norm. Essentially it is Tikhonov regularized least-squares or ridge regression. The problem becomes highly sensitive to random errors in the observed response y, producing a large variance. Small positive values of λ improve the conditioning of the problem and reduce the variance of the estimates. λ > 0 controls the trade-off between the fitting error and the size of x. It has the analytical solution for any λ > 0:

$$x = (A^T A + \lambda I)^{-1} A^T y. \tag{4.5}$$

The analytical solution has no restrictions on shape, rank of **A**.

Second, add the 1-norm penalty to the objective function and form the solvent for ℓ_1-regularized least squares problem:

$$\min_x \ \|Ax - y\|^2 + \lambda\|x\|_1 . \tag{4.6}$$

1-norm penalty introduces a sparse solution and constructs a convex heuristic method for finding a sparse x that gives **A**x ≈ y. Essentially it is lasso or basis pursuit and a non-differentiable problem due to norm-1 term. It does not have analytical solution but can be solved efficiently.

Third, add a combination between the ℓ_1 and ℓ_2 regularizations. Essentially it is an elastic net problem and seeks for solutions with highly correlated variables and lasso does not perform well. The solvent has the form:

$$\min_x \ \|Ax - y\|_2^2 \ + \ \lambda(\alpha\|x\|_2^2 + (1-\alpha)\|x\|_1) , \tag{4.7}$$

where $\alpha \in [0,1]$ and λ are parameters. When α = 0 it is lasso and when α = 1 it is a ridge regression. It is used when solutions are expected as groups of highly correlated variables (e.g. microarray, genes). It is a strictly convex problem for any α < 1 and λ > 0 (unique solution). Compared with

lasso, the elastic net considers the internal structure of vector x and selects the estimates of its components together. However, the lasso does not reflect the relative importance of its components.

Other types of constraints are used in least squares as well such as group lasso and fused lasso. Suitable regularization term should be based on the characteristics of the solution and prior information. Moreover, simulation can also be used to select the best regularization parameter and best regularization form.

4.4.2.Experiments: fluorophores separation

The mixed fluorophore pairs are Alexa Fluorophore 555 and Alexa Fluorophore 546, whose emission spectra are highly overlapped, but fluorescence lifetimes are different. Assume that 1), each fluorophore present has a characteristic lifetime mono-exponential decay; 2) Lifetime τ_1, τ_2 can be determined in advance following the forward model and noises are added to the time channel. Experiments are designed and performed for validation. Biological samples are fixed neurons and two kinds of cytoskeleton structures, Vimentin immunolabeled with Alexa Fluor 555 and Tubulin immunolabeled with Alexa Fluor 546, are distinguished by the lifetime contrast between the two fluorophores. Vimentin and Tubulin are both filament proteins and abundant in the cytoskeleton structure. As is presented in **Figure 4.5**, they are spectrally very similar but can be separated using the fluorescence lifetime information. Vimentin is represented in the green channel and Tubulin is represented in the blue channel.

To select the most reliable regularization form and regularization parameter, simulation is performed given that the ratio of concentrations of two fluorophores is 1:2 in the mixture. Two types of restoration error are defined: 1) Relative norm of the restoration error $\|x_\lambda - x^{(0)}\|/\|x^{(0)}\|$; 2) Ratio error $\left\|\frac{x_\lambda(2)}{x_\lambda(1)} - \frac{x^{(0)}(2)}{x^{(0)}(1)}\right\|$. Through simulation, ℓ_2 -regularization produces the most satisfactory results and the regularization parameter is 6.86 in the simulation. Given the structure of vector x in this simulation, x is not likely to be sparse since two fluorophores are co-localized and components of vector x are not highly correlated. Meanwhile, x should have relatively small values in the simulation given the normalized amplitude ratio. As a consequence, ℓ_2-regularization gives better construction results than ℓ_1-regularization and the combination of ℓ_2-regularization and ℓ_1-regularization.

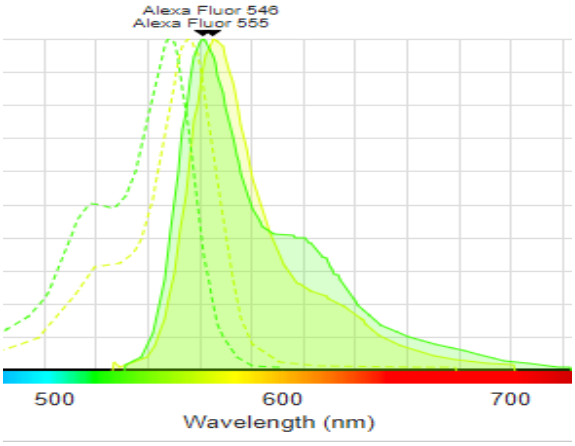

Figure 4.5 Emission spectra of Alexa Fluorophore 555 and Alexa Fluorophore 546

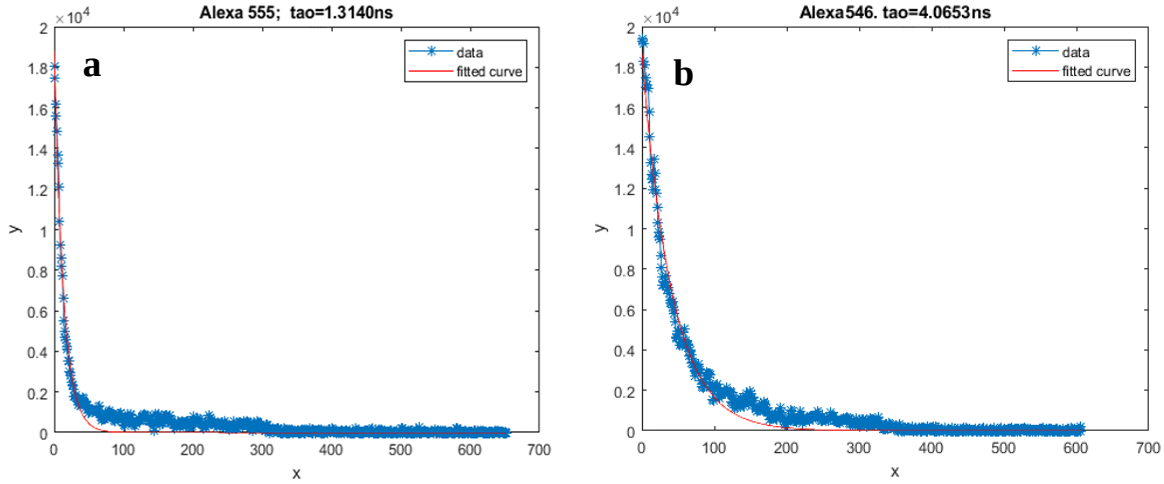

Figure 4.6 (a) Lifetime fitting of Alexa Fluorophore 555. **(b)** Lifetime fitting of Alexa Fluorophore 546. y axis represents the fluorescent intensity. x axis represents the time counting points.

The first procedure is a priori lifetime characterization of Alexa Fluorophore 555 and Alexa Fluorophore 546. The Alexa Fluorophore 555 and Alexa Fluorophore 546, which are in the form of conjugated secondary antibody, are directly imaged with the compressed FLIM. Time-domain images are acquired using the compressed FLIM with 513 nm laser excitation. The two lifetime components are fit from the asymptotic portion of the fluorescence decay curve. Through characterization illustrated in **Figure 4.6**, the two lifetimes are 1.3140 ns and 4.0653 ns, respectively.

The second procedure is the sample preparation. Fixed neurons are used for the staining procedure. Customized protocols are developed to adapt to the compressed FLIM system. Since fluorescence lifetime is sensitive to a wealth of environment factors such as pH, ion or oxygen

concentration, molecular binding or the proximity of energy acceptors, all the variables are controlled to be the same as the priori lifetime characterization procedure.

The third procedure is measurement of the stained biological samples with the compressed FLIM. All the system parameters are strictly controlled to be the same as the priori lifetime characterization procedure. The capture intensity image is presented in **Figure 4.7 a**.

The fourth procedure is unmixing from the measurement of the biological stained fixed neurons with the compressed FLIM system using the ℓ_2-regularization least squares algorithm. From the lifetime-based temporal unmixing, the amplitude ratio of Alexa Fluorophore 555 and Alexa Fluorophore 546 is recovered. Then the normalized amplitude ratio of Alexa Fluorophore 555 is assigned to the green channel of the RGB lifemap and the normalized amplitude ratio of Alexa Fluorophore 546 to the blue channel of the RGB lifemap. The recovered RGB lifemap image is presented in **Figure 4.7 b**.

Lifetime-based multiplexing enables simultaneous imaging of multiple fluorophores. This simple approach overcomes the spectral bandwidth limitations, and in contrast to frequency domain (FD), mono-exponential analysis, forward Laplace or moment-based approaches, determines fluorophore concentrations directly from the time-domain signal without the need for

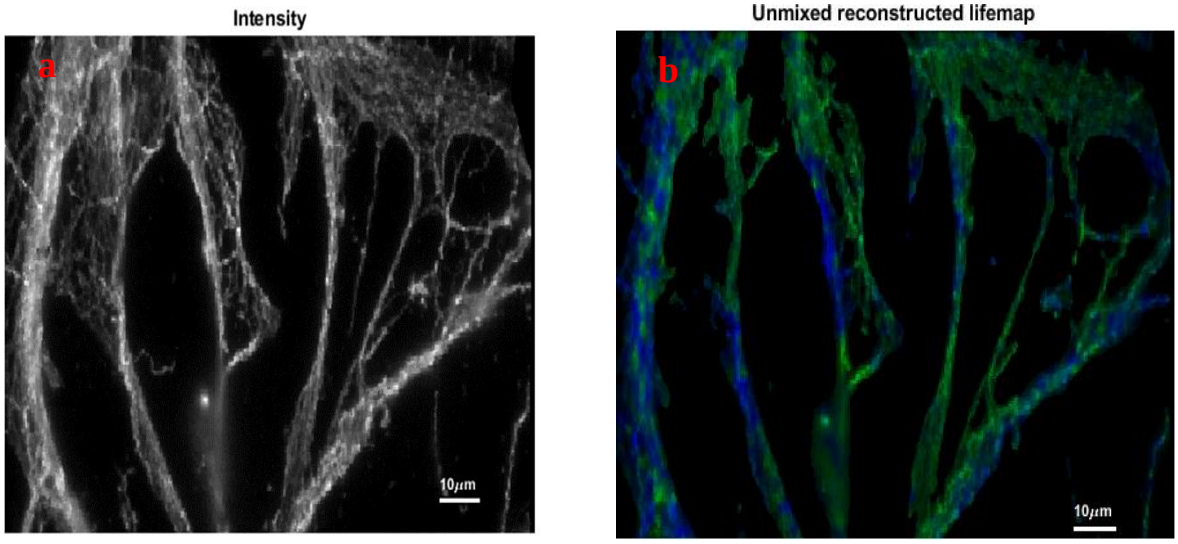

Figure 4.7 (a) Intensity image of stained fixed neurons with Alexa Fluorophore 555 and Alexa Fluorophore 546. (**b)** Unmixing lifemap of stained fixed neurons with Alexa Fluorophore 555 and Alexa Fluorophore 546. Blue channel represents the Tubulin and green channel represents the Vimentin.

additional data processing. Unmixed amplitudes are linear with fluorophore concentration and can be input directly into the lifemap rendering. Lifetime multiplexing should be generalizable to the full spectrum probe repertoire given a suitable detecting system, after proper a priori characterization of mixed fluorophore components decay characteristics. The fluorophores tested in this project have narrowly distributed, mono-exponential lifetimes and are therefore suitable for multiplexing. Moreover, this approach can be extended to in vivo lifetime multiplexing because it is expected that antibody-conjugated probes would exhibit similar in vivo lifetime characteristics.

Although mono-exponential probes are desirable, more complicated fluorophores could be imaged with multi-exponential fits (assuming priori lifetimes). In the case of a fluorophore with a distinct lifetime when accumulated at the target, two lifetimes could be used in a multiexponential fit to produce separate images of target-bound and unbound probes. With more complicated decay functions, such as autofluorescence, it may be possible to approximate fluorophore decays with multi or stretched exponential basis functions, although calculation of the lifetime-based unmixing problem in these cases may be more difficult. Lifetime-based fluorophores unmixing is a powerful technique that should allow simultaneously multi-labeled subcellular structure imaging and various biological applications. This lifetime-based unmixing algorithm permits quantifying the proportion of molecules within a volume within a microscopic region. Fluorophores can be designed to exhibit lifetime shifts upon target binding. With lifetime multiplexing, signals from the unbound probes due to imperfect uptake could be effectively removed, resulting in improved signal-to-noise reconstructions of the target-bound probes. Probes with two distinct lifetime states (ligand bound versus unbound) also offer the possibility of quantitative measurements of ligand concentration. For example, visible fluorescent Ca^{2+} sensors have been used to measure absolute calcium concentrations in solution and in vivo.

4.5. FLIM-FRET (Förster resonance energy transfer)

4.5.1. Förster resonance energy transfer (FRET)

Förster resonance energy transfer (FRET) consists of physical interactions of two fluorophores. Two requirements are needed for FRET occurrence. The first is close physical interaction between two fluorophores. Generally, the distance varies from 1 nm to 10 nm. FRET efficiency is in inverse proportion to the sixth power of the distance between the two fluorophores. The second is highly overlapped emission spectrum and absorption spectrum between two fluorophores. As a result, upon excitation, one fluorophore, named as the donor, will emit the fluorescence which can be

absorbed by another fluorophore, named as the acceptor. This functions as the excitation for the acceptor to further fluorescence. In some cases, the acceptor does not fluoresce. Overall, FRET results in the decreased fluorescence intensity and fluorescence lifetime of the donor.

4.5.2. Compressed FLIM of FRET in vitro

FRET involves the nonradiative transfer of excited state energy from a fluorophore, the donor, to another nearby absorbing but not necessarily fluorescent molecule, the acceptor. In the experiment, a pair consisting of a donor fluorophore (Alexa Fluor 546) and an acceptor fluorophore (Alexa Fluor 647) is examined. The emission spectrum of the donor overlaps considerably with the absorption spectrum of the acceptor, which is a requirement for a FRET pair. Acceptor bleed-through (ABT) contamination, i.e., the direct excitation of the acceptor at the donor's excitation wavelength, is avoided by the system excitation wavelength and emission filter as is demonstrated in the control experiment. Based on solutions containing a controlled amount of donor Alexa Fluor 546 and acceptor Alexa Fluor 647, fluorescence lifetime of the donor is analyzed by compressed FLIM. Three phosphate buffer solutions (PBS) are prepared with concentration ratio of donor Alexa Fluor 546 and acceptor Alexa Fluor 647 (A : D ratio) as tabulated in the bottom table in **Figure 4.8 c**. The sample solutions are injected into 3 glass tubes (14705-010; VWR). The upper part in the **Figure 4.8 c** is the three mixed solutions**.** Fluorescence intensity images are acquired using the reference camera in the compressed FLIM system. Fluorescence emission of the donor gradually diminishes with more acceptor presence and stronger fluorescence quenching. **Figure 4.8 b** shows the reconstructed fluorescence lifetime of the donor. The donor's lifetime gradually decreases as the acceptor's concentration increases. **Figure 4.8 a** shows the intensity image captured by the reference camera within the same field of view.

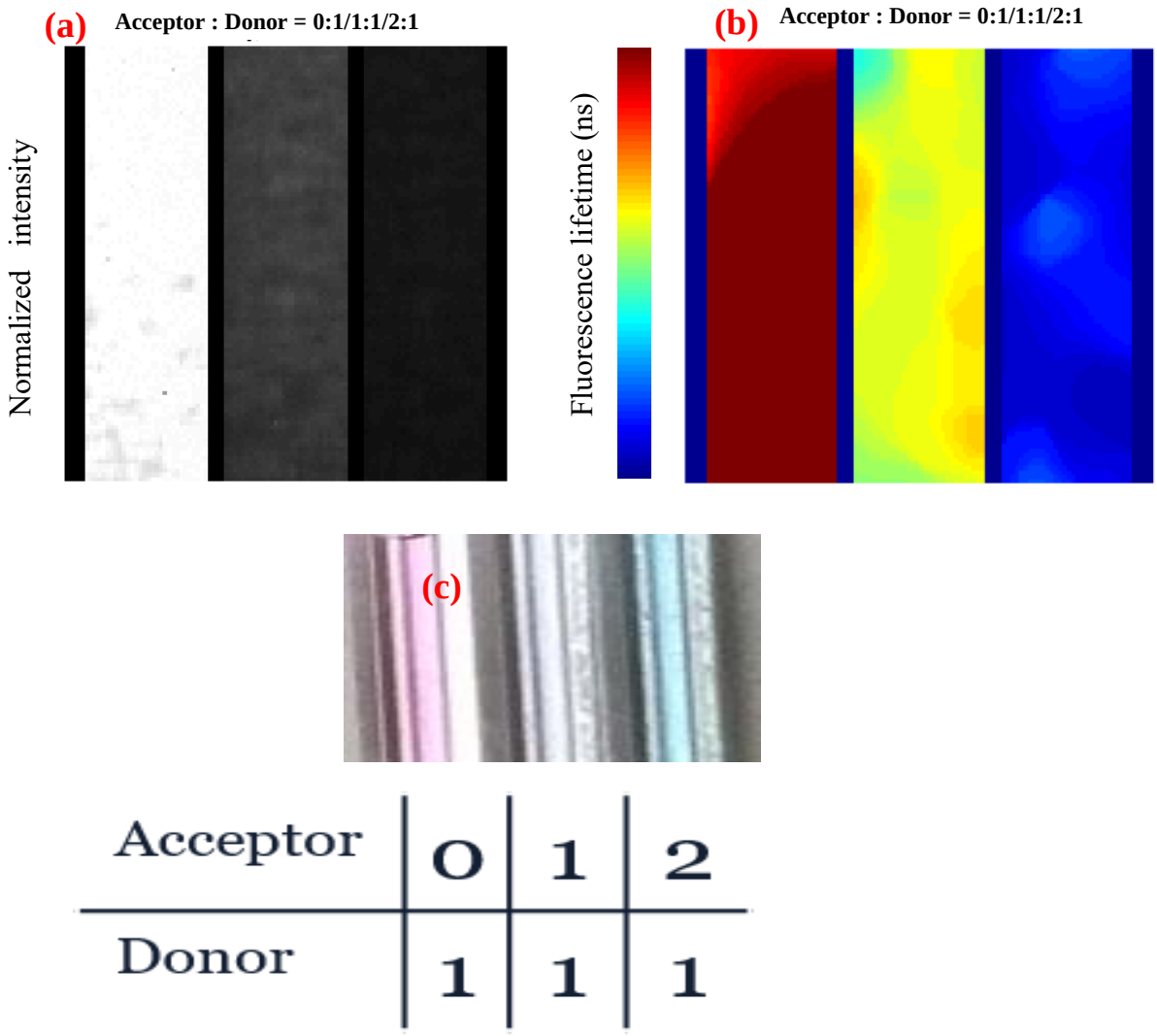

Figure 4.8 FRET imaging of compressed FLIM. Mixed solutions of different acceptor and donor concentration ratios are injected into three thin tubes and get imaged simultaneously by compressed FLIM. **(a)** Intensity image of three solutions. **(b)** Reconstructed lifetime image of the donors in the three solutions. **(c)** Three mixed solutions with different donor to acceptor concentration ratio.

4.6. Fluorescent beads diffusion dynamics

4.6.1.Lifetime contrast of fluorescent beads

Fluorescent beads of 6 µm diameter with orange spectral emission solutions are prepared and imaged. **Figure 4.9 a** presents reconstructed lifetime image of the single bead's solution by compressed FLIM. To further apply lifetime contrast to distinguish different kinds of beads with similar emission spectra, mixed beads solution containing 6 µm and 2 µm diameter fluorescent beads with similar orange fluorescent emission spectra are prepared. The two species can be distinguished based on their lifetime difference. **Figure 4.9 b** presents a reconstructed lifetime image of the mixed beads solution by compressed FLIM. **Figure 4.9 c** presents a reconstructed lifetime histogram of the mixed beads.

4.6.2.Beads diffusion dynamics

The diffusion behavior of single fluorescent beads is imaged. Fluorescent beads of 6 µm diameter with orange spectral emission solutions are prepared. Solvents are pipetted into the solutions slowly to introduce concentration gradient. The diffusion process is monitored by the compressed FLIM at 75 fps image acquisition speed. Each frame is color coded by the fluorescence lifetime of the beads using the compressed FLIM reconstruction. Further, mixed beads solution diffusion dynamics containing 6 µm and 2 µm diameter fluorescent beads with similar orange fluorescent emission spectra are imaged by compressed FLIM. The frame rate is restricted by the external reference camera which is 75 fps at its maximum in the external trigger mode.

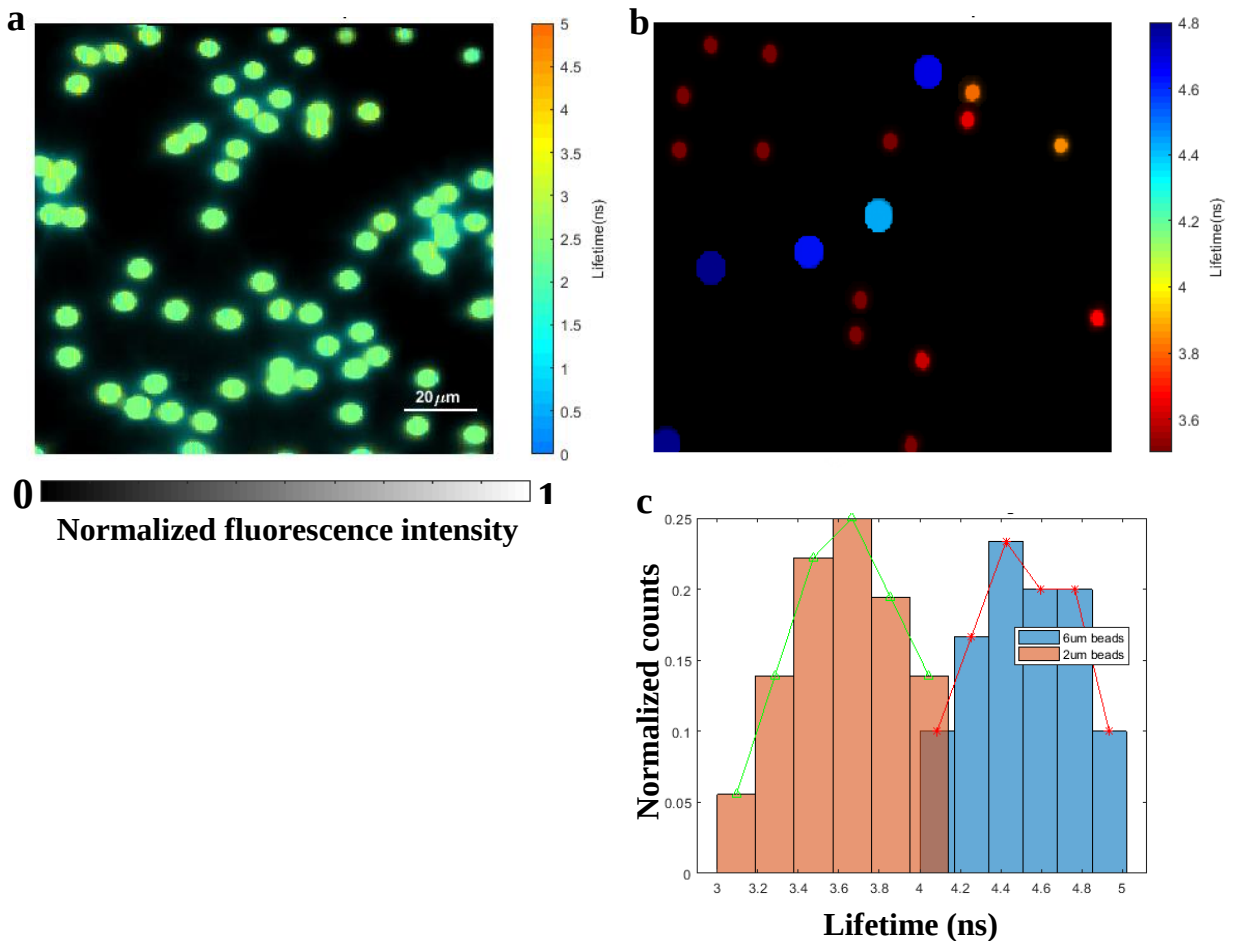

Figure 4.9 Compressed FLIM of fluorescent beads diffusion dynamics. (a) Reconstructed lifetime image of the single beads solution by compressed FLIM. **(b)** Reconstructed lifetime image of the mixed beads solution by compressed FLIM. **(c)** Reconstructed lifetime histogram of the mixed beads solution by compressed FLIM.

4.7. GPU-assisted real time construction of compressed FLIM

4.7.1. Graphics processing unit (GPU) acceleration

GPU is a highly parallel multi-processor optimized for 2D/3D graphics, video and display. GPU is the key part of image processing and the heart of computer graphics to handle complex geometric and mathematical calculations. By contrast, a central processing unit (CPU) has a few cores for sequential tasks. GPU has thousands of smaller cores for parallel tasks. Part of the original CPU's work is converted into GPU's work for higher computer working efficiency. More transistors are

used for data calculation than data extraction in GPU in contrast to CPU. They have similar purposes but are optimized for different computing tasks.

The GPU is specifically designed to solve data problems that can be expressed in parallel—a program where many data elements execute at the same time, with extremely high computational density. GPU executes very fast, but when it operates to get complicated data calculation results from memory, the entire pipeline has a long pause and the speed is slow. Inside the CPU, multi-level cache usage to increase the speed of accessing memory is conventional. As a consequence, CPU and GPU collaborate for higher computer efficiency in GPU acceleration.

4.7.2. GPU-assisted acceleration implementation

Under TwIST framework implementation of compressed FLIM, the operation codes are composed of several parts illustrated in the flow chart in **Figure 4.10**. In GPU acceleration, for the serial parts, the algorithm uses the computing power of the Illinois supercomputer. For the computer-intensive matrix operations in each function, the algorithm employs the GPU for the output. Functions including denoising, conv2c and projk are highly computationally intensive and are executed by the GPU for the output. For the communication among the functions and the objective function criteria, CPU is exploited for the output. To further speed up the reconstruction process and realize real-time compressed FLIM reconstruction, we use the hardware computing power of the Illinois campus computer cluster. The campus cluster file system hardware consists of 6 Dell R740 servers each with 32 cores/384 GB memory, dual enhanced data rate (EDR), fourteen data rate (FDR) and 40 GbE interface and 2 Samsung PM1725 3.2 TB NVME SSDs for metadata [36]. **Table 4.1** illustrates the reconstruction time comparison between original TwIST implementation and GPU-assisted TwIST implementation. Under the spatial resolution of 500 × 450 pixels, TwIST implementation takes around 3 minutes for one lifemap. With the GPU acceleration, we spend less than 3 seconds for one lifemap, a 57-fold improvement speedup.

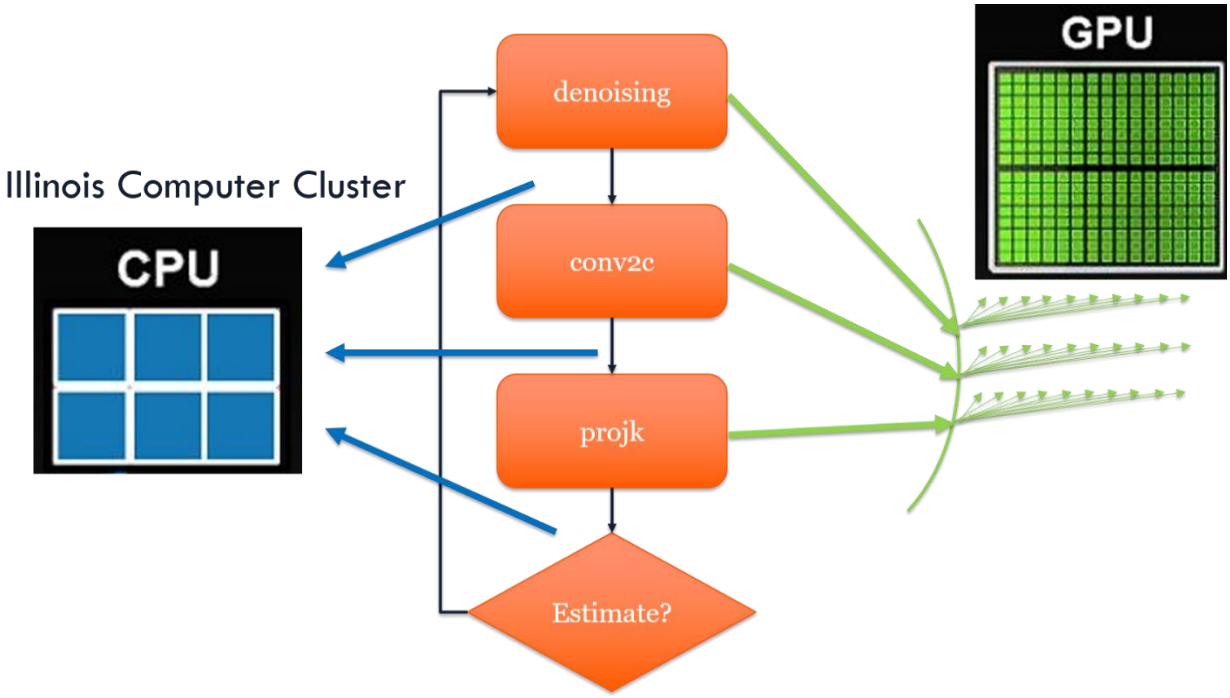

Figure 4.10 GPU-assisted real-time reconstruction of compressed FLIM under TwIST implementation. The flow chart presents compressed FLIM reconstruction main steps under TwIST framework. In GPU acceleration, for the serial parts, the algorithm uses the computing power of the Illinois supercomputer. For the computer-intensive matrix operations in each function, the algorithm employs the GPU for the output.

Table 4.1 Reconstruction time of conventional TwIST implementation and GPU-assisted TwIST implementation of 500 × 450 pixel lifetime image of compressed FLIM

500 × 450 pixels	**TwIST**	**GPU**
Reconstruction time	155s	2.7s
Speedup ratio	\	~57

5. Conclusions and future work

5.1. Conclusions

Compressed FLIM overcomes limitations of current FLIM systems by incorporating the paradigm of compressed sensing. It is built upon CUP technology and tailored for high-resolution and high-speed FLIM. It enables widefield, depth-resolved lifetime imaging at up to 100 fps, approximately two orders of magnitude faster than that allowed by current FLIM imagers. The processing speed for compressed FLIM lifetime image of 500 × 450 pixels size is 2.7 s after GPU-assisted acceleration. Compressed FLIM will have broad impacts on a variety of biomedical imaging applications, ranging from microscopy to endoscopy. Compressed FLIM will lead to a new generation of widefield lifetime imagers and enable high-speed recording of fluorescence lifetimes in transient biological events.

5.2. Neuron spiking dynamics imaging using FRET-opsin-based genetically encoded voltage indicators

Neurons are fundamental components of our nervous system responsible for communicating information throughout the body and enabling body memory, emotion, and abstract thought as well as basic reflexes. Inside a neuron, chemical signals from adjoining cells are converted to an electric action potential which propagates along the neuron's spindly axon until it reaches the end. Here at the synapse, the electrical signal is once again converted to the release of chemical signals, which via diffusion can relay the signal to the next nerve cell. Current therapies for neurological disorders rely on traditional medication and electric stimulation, are often encumbered by accessory equipment (i.e., tubes, pumps, valves) and have difficulty in determining precise dosage.

Action potentials in neurons, also known as spikes, are rapid rises and subsequent falls in membrane potential at the neuron axons. Neurons depend on the spike propagation along the axons toward the synaptic boutons for intercellular communications, and measuring the neural spikes is required to investigate neural activities. **Figure 5.1** presents typical neuron firing patterns in different kinds of neurons. Under strong depolarizing current, fast spiking (FS) cells fire at around 200 Hz, requiring a 200 Hz imaging speed. These cells could sustain high-frequency firing at 200–300 Hz in response to a strong depolarizing current. Regular spiking (RS) cells fire at less than 100 Hz. Intrinsically bursting (IB) cells fire from 250 Hz to 400 Hz and respond to a depolarizing current injection with single or multiple clusters, or bursts, of action potentials. Bursts consist of three to seven action potentials, with an average frequency within a burst ranging from 250 Hz to

400 Hz. Multiple bursts could often be evoked from 4 to 7 Hz in IB cells. FLIM provides a robust way for spiking detection because it can measure the absolute membrane voltage change of neuron spiking with fluorescent voltage sensors expressed in the neurons [37].

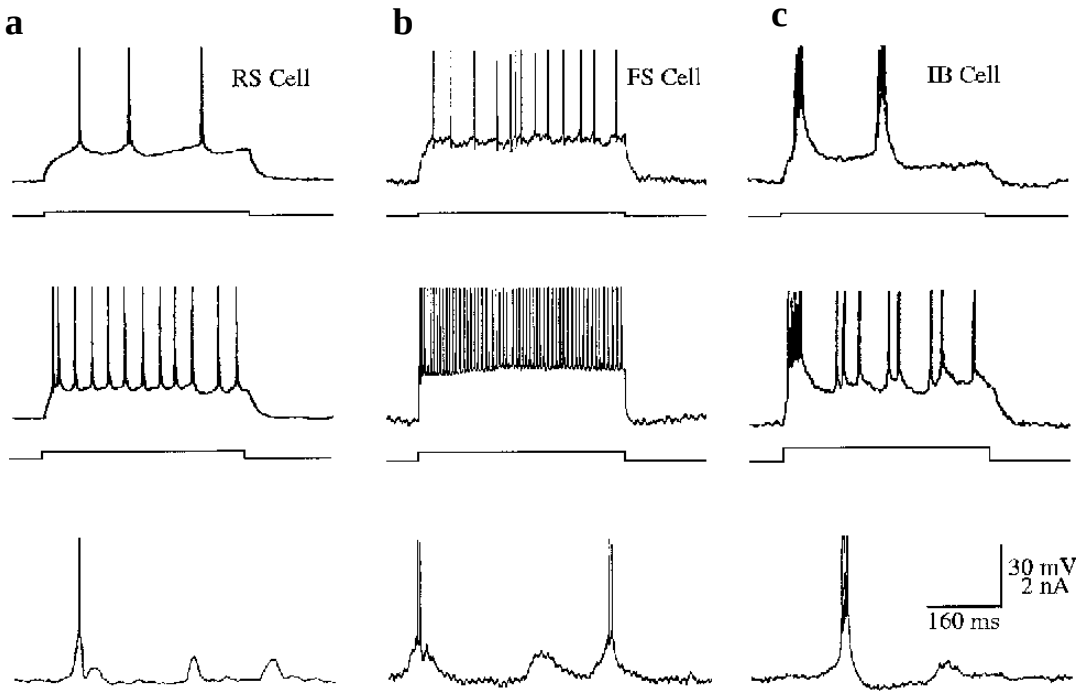

Figure 5.1 Neuron firing patterns under strong depolarization current. (a) Regular-spiking (RS) cells. **(b)** Fast-spiking (FS) cells. **(c)** Intrinsically bursting spiking (IB) cells [37]. Different kinds of neuron cells fire at different frequencies in response to a strong depolarizing current.

Genetically encoded voltage indicators (GEVIs) offer great promises in visualizing neural spiking dynamics [38]. Compared with traditional calcium sensors, GEVIs can read out the spike train with much faster kinetics, thus enabling optical recording of momentary neural activities. To image the dynamics of GEVIs, most methods rely on intensity-based measurements. In brief, the conformation or absorption of GEVIs is engineered such that they are sensitive to the transmembrane potential. During voltage depolarization, the induced conformation or absorption changes alter the brightness of GEVIs, which is then measured by an intensity camera. Despite high sensitivity, the intensity-based measurements infer only relative voltage changes, and the results are easily confounded by variations in excitation fluence, signal decay by photobleaching, and background fluorescence, particularly when the imaging is performed in vivo.

Compared with intensity-based approaches, FLIM provides a more quantitative assessment on the fluorescent probes because it is insensitive to aforementioned factors. This functional independency has recently enabled absolute voltage measurement in neurons. To test compressed FLIM in such an application, we will image a FRET-opsin-based GEVI, MacQ-mCitrine [39]. The sensor was constructed by fusing a bright fluorescent protein mCitrine, which serves as the FRET donor, to the Mac rhodopsin, which acts as both the voltage sensitive domain and FRET acceptor. During voltage depolarization, the rhodopsin's absorption increases. Because rhodopsin and mCitrine are within the FRET-effective distance, the increase of rhodopsin's absorption affects mCitrine's excitation state by changing the branching ratio between fluorescence and nonradiative decay, a fact which reduces mCitrine's fluorescence emission as well as its lifetime. Compared with monolayer cultured cells, previous studies showed that the apparent sensitivity of intensity-based approaches degrades in brain slices because of tissue autofluorescence and out-of-focus fluorescence [39, 40]. Because of optical sectioning, we expect compressed FLIM can outperform conventional intensity-based methods in this gold-standard model system.

5.3. Protein folding kinetics imaging using temperature-jump relaxation method

Proteins are made up of specified sequences of amino acid and are necessary to our fitness. Amino acid chains have to fold into the proper 3D structure before they can execute the protein's specific function. The protein folding mechanism is extensively studied. A peptide can fold into its native structure within an extremely short period of time down to a fraction of a microsecond [41]. Extensive computational and experimental efforts have been devoted to determining the kinetics of protein folding, revealing that the transient events on the folding pathway are vital to our understanding of the molecular mechanisms of proteopathies, such as Alzheimer's disease and Parkinson's disease. Particularly, fast-folding proteins have been a major focus of computational and experimental study because they are accessible to both techniques. From the computational perspective, small proteins with a short polypeptide chain reduce the configurational search space and therefore can be reasonably simulated with current computational power, which has just recently reached from the nanosecond to the microsecond domain [42]. **Figure 5.2** represents a typical simulated fast protein folding process by molecular dynamics (MD) simulation from native state at the start to the folded state after 2500 ns. The most representative structures from clustering analysis are illustrated with the corresponding time stamps [43]. Meanwhile, from the experimental perspective, small proteins avoid imperfection intrinsic to the folding process—the

longer the protein sequence, the more likely there are to be unfavorable interactions, which cannot be simultaneously eliminated by protein evolution or engineering. Therefore, small proteins—with their minimally frustrated folding landscapes—fold fast, and they can be re-engineered to fold even faster [44], approaching the theoretical speed limit for folding [44]. However, despite the vast availability of small proteins that are amenable to MD simulation, their folding times cluster in ns-μs, a range into which very few experimental observation techniques can peek. Effectively resolving this event requires a GHz temporal sampling rate, which is beyond the bandwidth of most electronic sensors.

Currently, the primary experimental tools to study protein folding encompass stopped flow, nuclear magnetic resonance (NMR) spectroscopy, and relaxation methods [45]. Conventional stopped flow and NMR spectroscopy have a relatively low temporal resolution (sub-millisecond), making them too slow for the study of fast folders. By contrast, relaxation methods, adapted from techniques developed in the 1950s–1970s to study chemical reactions in real time [46], enable measurement of folding events occurring at a much faster timescale.

Relaxation methods are based on the application of a sudden small perturbation (in temperature or pressure, for example) to study the recovery of equilibrium. To make a relaxation experiment work for fast-folding proteins, it requires two key ingredients: a fast initiation step and a fast detection scheme. Among various fast initiation techniques, laser temperature jumps are the most commonly used method because of their extremely short dead times (down to 2 ps). In practice, most laser temperature-jump experiments are carried out with nanosecond heating pulses to pump directly H_2O or D_2O stretching vibrations, which thermalize within the nanosecond pump profile. On the other hand, to resolve the relaxation of fast folders, the detection system must provide high temporal resolution in a nanosecond scale, which is usually achieved with a pump-probe-based method. However, the pump-probe-based methods require the folding itself to be strictly repeatable after different time delays after initialization of relaxation. Since folding transition is inherently a stochastic process, the pump-probe-based methods present difficulties in resolving the burst phase of the folding/unfolding, and they are inapplicable to single-molecule folding experiments. Small protein folding times cluster from microseconds to nanoseconds, and effectively resolving this event requires a GHz temporal sampling rate, which is beyond the bandwidth of most electronic sensors. FLIM relies on repetitive measurement, the laser repetition rate is around 100 MHz and it requires the folding to be strictly repeatable in each excitation cycle

and delay measurement. However, folding transition is inherently a stochastic process and requires snapshot measurement, so it is hard for the repetitive methods to resolve the protein folding.

To apply compressed FLIM for protein folding studies, we will integrate it with an imaging-based temperature-jump relaxation method. The spatially and temporally resolved folding dynamics after perturbation can then be inferred from the compressed FLIM measurement of fluorescence signals of endogenous chromophores such as tryptophan or the exogenous biosensors such as FRET. Slow protein folding will be investigated through FRET. During the folding process, if the protein is labelled with FRET pairs, the structure dynamics will result in a change of the distance between the FRET pairs, which changes the fluorescence lifetime of the donor. Using the compressed FLIM to detect the FRET in the folding process will enable inferring the structure change of the protein. Based on the slow protein folding result, fast protein detection can be explored and developed.

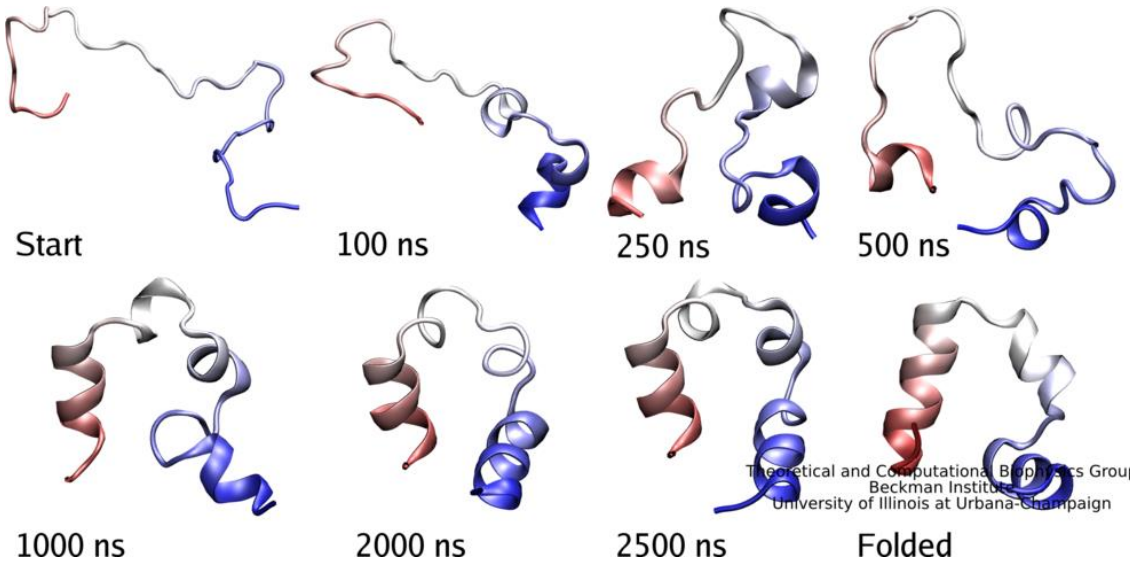

Figure 5.2 Simulated protein folding process by molecular dynamics (MD) simulation from native state at the start to the folded state after 2500 ns. The most representative structures from clustering analysis are illustrated with the corresponding time stamps [43].

5.4. Machine learning for compressed FLIM reconstruction

Machine learning methods have developed rapidly in the past decade with the development of computer hardware and software, and can be applied to solve the inverse problem in image reconstruction. These methods feature high efficiency and high speed, and offer the potential to solve complicated problems. The techniques used are different from the model-based method. The aim is to design a deep neural network (DNN), which is a special mapping function that directly maps the input to the output without time-consuming iterations and optimization process. The key is the parameters, which are learned through the large dataset. Nguyen et al. [47] developed a convolutional neural network based on Fourier ptychographic microscopy to reconstruct time sequences of dynamic live cells. Their technique images the cells efficiently and precisely in real time. It is impossible for a model-based method to finish this task because the cell environment is complicated and unpredictable. A model-based method cannot extract an effective mathematical abstract for this inverse problem and the time cost is unaffordable for real-time imaging. Furthermore, Tian and Waller [48] discussed the application of machine learning to 3D microscopy and proved its efficacy. For 3D microscopy, with pre-designed parameters, machine-learning-methods can accomplish the task in real time and with high accuracy, indicating that machine-learning methods can attack high-dimensional data sets and can be applied to more intricate tasks in the future.

On the other hand, without a well-designed DNN, machine-learning can fall short. Machine-learning cannot give convincing mathematical expressions and lacks solid theoretical grounding. Lucas et al. [49] reported that machine learning can give ridiculous results due to the lack of a precise analytical model. Given the input, it produces unreliable outputs that are much worse than what the model-based method gives. Tian and Waller [48] discussed the large training data and high time cost for machine learning methods. To get the well-defined parameters for the network, a lengthy training process and extra-large training data are required; generally, it costs from hours to several days to get a set of parameters and an optimization process may also be needed.

Based on the forward response of the compressed FLIM, we can construct a deep learning network and train the network with simulated samples and experimental samples. We can also combine the network with traditional model-based algorithms, then compare the reconstruction results to improve the reconstruction quality.

References

[1] L. V. Wang and H.-I. Wu, Biomedical optics: principles and imaging, Hoboken: John Wiley & Sons, 2012.

[2] M. T. Spence and I. D. Johnson, The molecular probes handbook: A guide to fluorescent probes and labeling technologies, Carlsbad: Live Technologies Corporation, 2010.

[3] J. R. Lakowicz, Principles of fluorescence spectroscopy, Springer Science & Business Media, 2013.

[4] J. W. Borst and J. V. Antonie, "Fluorescence lifetime imaging microscopy in life sciences," Measurement Science and Technology, vol. 21, no. 10, p. 102002, 2010.

[5] P. J. Verveer, F. S. Wouters, A. R. Reynolds and P. I. H. Bastiaens, "Quantitative imaging of lateral ErbB1 receptor signal propagation in the plasma membrane," Science, vol. 290, no. 5496, pp. 1567-1570, 2000.

[6] C. Chang, D. Sud and M. Mycek, "Fluorescence lifetime imaging microscopy," Methods in Cell Biology, vol. 81, pp. 495-524, 2007.

[7] W. Becker, "DCS-120 confocal scanning FILM systems," 2015. [Online]. Available: http://www.tokyoinst.co.jp/product_file/file/BH03_tec01_ja.pdf.

[8] R. V. Krishnan, A. Masuda, V. E. Centonze and B. Herman, "Quantitative imaging of protein–protein interactions by multiphoton fluorescence lifetime imaging microscopy using a streak camera," Journal of Biomedical Optics, vol. 8, no. 3, pp. 362-367, 2003.

[9] A. J. Bower, J. Li, E. J. Chaney, M. Marjanovic, D. R. Spillman and S. A. Boppart, "High-speed imaging of transient metabolic dynamics using two-photon fluorescence lifetime imaging microscopy," Optica, vol. 5, no. 10, p. 1290-1296, 2018.

[10] "Guide to streak cameras," Photonics, Hamamatsu, Hamamatsu City, Japan, 2008.

[11] Hamamatsu Photonics, "Guide to streak cameras," 2008. [Online]. Available: https://www.hamamatsu.com/resources/pdf/sys/SHSS0006E_STREAK.pdf.

[12] L. Gao, J. Liang, C. Li and L. V. Wang, "Single-shot compressed ultrafast photography at one hundred billion frames per second," Nature, vol. 516, no. 7529, pp. 74-77, 2014.

[13] J. Requejo-Isidro, J. McGinty, I. Munro, D. Elson, N. Galletly, M. Lever, M. Neil, G. Stamp, P. French, P. Kellett and J. Hares, "High-speed wide-field time-gated endoscopic fluorescence-lifetime imaging," Optics Letters, vol. 29, no. 19, pp. 2249-2251, 2004.

[14] D. Elson, I. Munro, J. Requejo-Isidro, J. McGinty, C. Dunsby, N. Galletly, G. Stamp, M. Neil, M. Lever, P. Kellett and A. Dymoke-Bradshaw, "Real-time time-domain fluorescence lifetime imaging including single-shot acquisition with a segmented optical image intensifier," New Journal of Physics, vol. 6, no. 1, p. 180, 2004.

[15] D. Grant, J. M. E. McGinty, T. Bunney, D. Owen, C. Talbot, W. Zhang, S. Kumar, I. Munro, P. Lanigan and G. Kennedy, "High speed optically sectioned fluorescence lifetime imaging permits study of live cell signaling events," Optics Express, vol. 15, no. 24, pp. 15656-15673, 2007.

[16] A. Agronskaia, L. Tertoolen and H. Gerritsen, "High frame rate fluorescence lifetime imaging," Journal of Physics D: Applied Physics, vol. 36, no. 14, p. 1655, 2003.

[17] A. V. Agronskaia, L. Tertoolen and H. C. Gerritsen, "Fast fluorescence lifetime imaging of calcium in living cells," Journal of Biomedical Optics, vol. 9, no. 6, pp. 1230-1238, 2004.

[18] D. Li, D. W. R. Tyndall, J. Richardson, R. Henderson, J. S. D. Arlt and E. Charbon, "Video-rate fluorescence lifetime imaging camera with CMOS single-photon avalanche diode arrays and high-speed imaging algorithm," Journal of Biomedical Optics, vol. 16, no. 9, p. 096012, 2011.

[19] W. Yan, X. Peng, J. Qi, J. Gao, S. Fan, Q. Wang, J. Qu and H. Niu, "Dynamic fluorescence lifetime imaging based on acousto-optic deflector," Journal of Biomedical Optics, vol. 19, no. 11, p. 116004, 2014.

[20] S. Moon, Y. Won and D. Kim, "Analog mean-delay method for high-speed fluorescence lifetime measurement," Optics Express, vol. 17, no. 4, pp. 2834-2849, 2009.

[21] Y. Won, S. Moon, W. Yang, D. Kim, W. Han and D. Kim, "High-speed confocal fluorescence lifetime imaging microscopy (FLIM) with the analog mean delay (AMD) method," Optics Express, vol. 19, no. 4, pp. 3396-3405, 2011.

[22] A. Tsikouras, Q. Fang and A. Yeh, "Instrumentation design of a high-speed fluorescence lifetime imaging microscope tailored to high-throughput screening for drug discovery," in ECS Meeting Abstracts, Orlando, 2014.

[23] J. Bioucas-Dias and M. Figueiredo, "A new TwIST: Two-step iterative shrinkage/thresholding algorithms for image restoration," IEEE Transactions on Image Processing, vol. 16, no. 12, pp. 2992-3004, 2007.

[24] T. Niehörster, A. Löschberger, I. Gregor, B. Krämer, H. Rahn, M. Patting, F. Koberling, J. Enderlein and M. Sauer, "Multi-target spectrally resolved fluorescence lifetime imaging microscopy," Nature Methods, vol. 13, no. 3, p. 257, 2016.

[25] I. Gregor and M. Patting, "Pattern-based linear unmixing for efficient and reliable analysis of multicomponent TCSPC data," in Advanced Photon Counting, Springer, 2014, pp. 241-263.

[26] C. Wang, R. Xu, W. Tian, X. Jiang, Z. Cui, M. Wang, H. Sun, K. Fang and N. Gu, "Determining intracellular temperature at single-cell level by a novel thermocouple method," Cell Research, vol. 21, no. 10, pp. 1517-1519, 2011.

[27] K. Goda, K. Tsia and B. Jalali, "Serial time-encoded amplified imaging for real-time observation of fast dynamic phenomena," Nature, vol. 458, pp. 1145-1149, 2009.

[28] K. Nakagawa, A. Iwasaki, Y. Oishi, R. Horisaki, A. Tsukamoto, A. Nakamura, K. Hirosawa, H. Liao, T. Ushida, K. Goda and F. Kannari, "Sequentially timed all-optical mapping photography (STAMP)," Nature Photonics, vol. 8, no. 9, pp. 695-700, 2014.

[29] N. Matlis, A. Axley and W. Leemans, "Single-shot ultrafast tomographic imaging by spectral multiplexing," Nature Communications, vol. 3, p. 1111, 2012.

[30] B. Heshmat, G. Satat, C. Barsi and R. Raskar, "Single-shot ultrafast imaging using parallax-free alignment with a tilted lenslet array," in CLEO: Science and Innovations, San Jose, 2014.

[31] H. Shiraga, M. Nakasuji, M. Heya and N. Miyanaga, "Two-dimensional sampling-image x-ray streak camera for ultrafast imaging of inertial confinement fusion plasmas," Review of Scientific Instruments, vol. 70, no. 1, pp. 620-623, 1999.

[32] L. Wang, Z. Xiong, D. Gao, G. Shi and F. Wu, "Dual-camera design for coded aperture snapshot spectral imaging," Applied Optics, vol. 54, no. 4, p. 848–858, 2015.

[33] Y. Liu, X. Yuan, J. Suo, D. Brady and Q. Dai, "Rank minimization for snapshot compressive imaging," IEEE Transactions on Pattern Analysis and Machine Intelligence, p. 1, 2018.

[34] X. Liao, H. Li, and L. Carin, “Generalized alternating projection for weighted minimization with applications to model-based compressive sensing,” SIAM Journal on Imaging Sciences, vol. 7, no. 2, pp. 797–823, 2014.

[35] S. Raymond, D. Boas, B. Bacskai and A. Kumar, "Lifetime-based tomographic multiplexing," Journal of Biomedical Optics, vol. 15, no. 4, p. 046011, 2010.

[36] "Illinois Campus Cluster Program," Illinois Technology Services, 2018. [Online]. Available: https://campuscluster.illinois.edu/about/system-info/storage-environment/.

[37] D. Dombeck, L. Sacconi, M. Blanchard-Desce and W. Webb, "Optical recording of fast neuronal membrane potential transients in acute mammalian brain slices by second-harmonic generation microscopy," Journal of Neurophysiology, vol. 94, no. 5, pp. 3628-3636, 2005.

[38] Y. Gong, "The evolving capabilities of rhodopsin-based genetically encoded voltage indicators," Current Opinion in Chemical Biology, vol. 27, p. 84–89, 2015.

[39] D. Brinks, A. Klein and A. Cohen, "Two-photon lifetime imaging of voltage indicating proteins as a probe of absolute membrane voltage," Biophysical Journal, vol. 109, no. 5, p. 914–921, 2015.

[40] Y. Gong, M. Wagner, J. Li and M. Schnitzer, "Imaging neural spiking in brain tissue using FRET-opsin protein voltage sensors," Nature Communications, vol. 5, p. 3674, 2014.

[41] C. Dobson, "Protein misfolding, evolution and disease," Trends in Biochemical Sciences, vol. 24, no. 9, pp. 329-332, 1999.

[42] J. Shea and C. Brooks III, "From folding theories to folding proteins: A review and assessment of simulation studies of protein folding and unfolding," Annual Review of Physical Chemistry, vol. 52, no. 1, pp. 499-535, 2001.

[43] P. Freddolino, F. Liu, M. Gruebele and K. Schulten, "Ten-microsecond molecular dynamics simulation of a fast-folding WW domain," Biophysical Journal, vol. 94, no. 10, pp. L75-L77, 2008.

[44] C. Snow, B. Zagrovic and V. Pande, "The Trp Cage: folding kinetics and unfolded state topology via molecular dynamics simulations," Journal of the American Chemical Society, vol. 124, no. 49, p. 14548–14549 , 2002.

[45] M. Gruebele, "Fast relaxation methods," in Protein Folding Handbook, Weinheim, WILEY-VCH Verlag GmbH & Co. KGaA, 2005, pp. 454-490.

[46] M. Eigen and L. D. Maeyer, "Relaxation methods," in Techniques in Organic Chemistry, Part II , New York, Interscience, 1963, pp. 895-1054.

[47] T. Nguyen, Y. Xue, W. Tahir, Y. Li, L. Tian and G. Nehmetallah, "A deep-learning approach for high-speed Fourier ptychographic microscopy," in Imaging Systems and Applications, Orlando, 2018.

[48] L. Tian and L. Waller, "3D intensity and phase imaging from light field measurements in an LED array microscope," Optica, vol. 2, no. 2, pp. 104-111, 2015.

[49] A. Lucas, M. Iliadis, R. Molina and A. Katsaggelos, "Using deep neural networks for inverse problems in imaging: Beyond analytical methods," IEEE Signal Processing Magazine, vol. 35, no. 1, pp. 20-36, 2018.